Low-volume coronary artery bypass surgery: Measuring and optimizing performance

David M. Shahian, MD,^a and Sharon-Lise T. Normand, PhD^b

See related article on page 1306.

In the current issue of the *Journal*, Miyata and colleagues¹ describe the relationship between coronary artery bypass grafting (CABG) procedural volume and outcome in Japan. In reality, however, there are no high-volume programs in this study. What the authors have actually provided us is the most extensive study of low-volume and extremely low-volume CABG surgery in the literature. It complements previous studies from the United States that include some programs with low volumes, and it provides a striking counterpoint to New York studies that are weighted toward the high end of the volume spectrum.

This report illustrates the potential for good performance at low volumes, as well as the statistical challenge of accurately measuring performance when sample sizes are small. It raises a number of unresolved issues in the ongoing volume–outcome debate, at least as applied to CABG surgery. For example, some payers and other stakeholders continue to promote best-practice volume requirements that are increasingly beyond the grasp of many programs, particularly as overall CABG volumes decrease nationally. Is this appropriate policy given the available outcomes data? Because many lower-volume programs function at a high level, can the public be protected while at the same time not penalizing such excellent programs? Is there a rational lower volume limit for CABG surgery programs? Are there better ways to measure performance that are less compromised by small sample sizes? Are there specific process and structural approaches that might promote optimal functioning of small programs?

Is Low-volume CABG Surgery a Performance Problem or a Measurement Problem?

CABG is unique: it is a mature, standardized procedure that is performed more frequently than any other complex operation and that has also been scrutinized more thoroughly. Notwithstanding the general validity of the volume–outcome relationship for a number of medical conditions and surgical procedures, data from a variety of sources suggest that many low-volume CABG providers achieve excellent results.

We believe the fundamental issue with low-volume CABG surgery is not inherently poor performance but rather the difficulty in accurately measuring performance. These 2 perspectives have quite different implications. If it were clear that low-volume CABG providers were uniformly poor performers, immutably limited by their lack of sufficient “practice,” then the only reasonable solution would be volume thresholds. For some highly complex but very infrequently performed procedures, such as esophagectomy or pancreatectomy, this might well be a justifiable approach.

In reality, however, excellent performance is achieved by many CABG providers whose volumes, although they might number in the hundreds annually, do not meet the thresholds of organizations like the Leapfrog Group (450 procedures per year). In this circumstance, volume standards would unfairly stigmatize or penalize such high-quality but low-volume providers. Furthermore, at a time when CABG volumes

From the Department of Surgery, Center for Quality and Safety, and the Institute for Health Policy, Massachusetts General Hospital, and Harvard Medical School,^a and the Department of Health Care Policy, Harvard Medical School, and the Department of Biostatistics, Harvard School of Public Health,^b Boston, Mass.

Received for publication Nov 9, 2007; accepted for publication Dec 18, 2007.

Address for reprints: David M. Shahian, MD, Center for Quality and Safety, Massachusetts General Hospital, 55 Fruit St, Boston, MA 02114 (E-mail: dshahian@partners.org).

J Thorac Cardiovasc Surg 2008;135:1202-9
0022-5223/\$34.00

Copyright © 2008 by The American Association for Thoracic Surgery

doi:10.1016/j.jtcvs.2007.12.037

are decreasing, such thresholds could also have unintended negative consequences. Given the importance of cardiac surgery to most institutions, failure to meet guidelines for “center of excellence” status or premium reimbursement might have substantial adverse implications. This could result in a perverse incentive to relax appropriateness criteria to meet volume thresholds, which might have a net negative effect on the health system.

If one views low-volume CABG providers as a heterogeneous group, many of whom provide excellent results, then the main issue is how to accurately measure the performance of individual programs, a challenge with small sample sizes and limited mortality events. More comprehensive and robust approaches to performance measurement could be developed that are less limited by such concerns, and specific programmatic initiatives could also be implemented to facilitate high performance in smaller programs.

We will examine both the evidence for a CABG volume–outcome relationship as well as statistical problems with assessing performance in low-volume programs. Findings from the study of Miyata and colleagues¹ will be reviewed in the context of these two issues. Finally, recommendations will be presented to enhance both performance and its measurement in low-volume CABG programs.

Previous CABG Volume–outcome Studies

Although the strength of the CABG volume–outcome relationship is probably exaggerated in some studies by failure to account for sample size and clustering (eg, through the use of hierarchical models),² there is little question that some association exists. This was evident in the original work of Luft and associates³ nearly 30 years ago, and it has been demonstrated in numerous subsequent studies, including those from the modern era.^{4–11} The strongest data supporting a volume–outcome association come from New York, although these studies include very few programs that are truly low volume, and their findings might not be generalizable. In 2004, for example, there were 39 New York programs providing isolated CABG surgery, and 75% of these programs had volumes of greater than 214 procedures. It remains uncertain whether the CABG volume–outcome relationship applies to all patients¹⁰ or primarily to those at higher risk.^{12,13}

Notably, the volume strata and mortality ranges for CABG are quantitatively unique among complex procedures in which the volume–outcome association has been investigated. In studies by Birkmeyer and associates⁴ using claims data on 901,667 Medicare patients, low and high volume hospital categories for CABG were < 230 procedures and > 849 procedures respectively, with an absolute mortality difference of only 1.1% (4.5% vs 5.6%) between these two extremes. In contrast, the low and high volume ranges were <1 and >16 procedures for pancreatectomy and <2 and >19 procedures for esophagectomy, with absolute differences in adjusted mortality that were orders of magnitude greater (16.3% vs

3.8% for pancreatectomy, 20.3% vs 8.4% for esophagectomy) than those for CABG.

Perhaps because it is a mature and frequently performed procedure, the volume–outcome association for CABG is weak. Studies by Peterson and coworkers¹³ using 2000–2001 data from the Society of Thoracic Surgeons (STS) National Adult Cardiac Database, adjusted for risk factors and clustering, demonstrated only a 0.07% decrease in mortality for every additional 100 procedures ($P = .004$). Because there was substantial variability in mortality in all strata of volume, there was limited ability to discriminate among providers based solely on volume. Similar findings were noted in a study of 228,738 patients by Rathore and associates⁷ using data from the 1998–2000 Nationwide Inpatient Sample. In both studies, the vast majority of low-volume providers are distributed widely and symmetrically about the mean, with variation increasing at progressively smaller program volumes. Scatterplots of observed mortality versus volume in these studies look strikingly similar to the funnel appearance of the 95% confidence intervals of a binomial event, with an average occurrence rate of about 2% to 3%, taken at various sample sizes.¹⁴ This is illustrated in Figure 1, a scatterplot based on the 2004 isolated CABG results from Massachusetts, New York, Ontario, and California. Superimposed scatterplot smoothers are roughly horizontal, showing little volume–outcome association. Much of the variability in mortality at low volumes, regarded by many as an indicator of inconsistent performance, is quite likely explained by sampling error.

Outcomes Profiling in Low-volume Programs

An alternative approach to volume thresholds is outcomes profiling. Public reporting of CABG outcomes is favored by many policymakers and has been mandated by law in states like New York and Massachusetts. Properly performed (by no means a trivial caveat), such reports are reasonably objective, they provide transparency and accountability, and they address the most important interest of patients: operative survival. However, they are the most demanding in terms of the need for high-quality data, audit and validation, and appropriate analytic methodologies. Even with larger sample sizes, comparative assessment of provider performance can be challenging, especially when based on a single outcome such as mortality. This becomes increasingly problematic as sample sizes (program volumes) decrease, a feature illustrated previously with regard to volume–outcome studies.

Outcomes profiling generates estimates of provider performance derived from a snapshot in time, typically a year of clinical activity. Such observed results are used to estimate “true” underlying program quality, ideally with confidence intervals that indicate how certain we are about this point estimate. As noted previously, the statistical confidence intervals around point estimates of mortality, an infrequent binomial event, become quite wide with small sample sizes (annual program volumes).⁵ Much of the variation of annual

Figure 1. Scatterplot of 2004 coronary artery bypass grafting (CABG) volume (*x*-axis) and risk-adjusted mortality rates (RAMR, *y*-axis) from New York, Massachusetts, California, and Ontario. Superimposed scatterplot smoothers are relatively flat. There is wide variability in outcomes at low volume, which is predominately sampling error. In the 2001–2004 Japanese CABG experience reported by Miyata and colleagues,¹ approximately 94% of programs would lie to the left of the vertical line drawn at a volume of 100 procedures annually (estimated by averaging the results over these 4 years). Individual hospital data from Japan were not available.

mortality rates among low-volume programs (eg, fewer than 100–150 procedures per year), as shown in Figure 1, can be largely explained by random statistical fluctuation, and this in turn limits the ability to draw firm conclusions about program quality.^{5,15} In studies by Dimick and colleagues,¹⁵ CABG was the only complex procedure performed with sufficient frequency by most programs to detect a doubling of mortality rate based on 3-year aggregate data. However, as CABG mortality rates continue to decrease, the sample sizes necessary to detect meaningful differences increase correspondingly.

The Japanese Experience: Extremely Low-volume CABG Surgery

This brings us to the study by Miyata and colleagues¹ in the current issue of the *Journal*, an extreme and revealing illustration of both the “problems” of low-volume surgery, as well as some potential solutions. The authors describe the demographics of CABG programs in Japan, where annual volumes are so uniformly low that there is simply no US analog.^{3,4,7} In Japan there has been limited regulatory oversight of cardiac surgery program proliferation or performance prior to the past few years. Relative to their population size, lower incidence of coronary disease, and number of isolated CABG procedures, the number of CABG providers in Japan far exceeds that of any publicly reported US states or the province of Ontario, as demonstrated in Table 1. Overall, 76% of Japanese programs perform fewer than 50 CABG procedures annually. Only 5.6% of 540 Japanese cardiac surgery programs (representing most of the programs in the country) performed at least 100 CABG procedures annually between January 2001 and December 2004, and 24.6% of programs performed fewer than 15 procedures annually. Median an-

nual CABG volume was 28 procedures per year during this period (interquartile range, 15–49 procedures), and the average annual volumes ranged from 0.25 to 293 isolated CABG procedures. Using a threshold of 150 procedures annually, 98.3% of Japanese programs would be classified as low or very low volume by US standards. Based on the findings of Miyata and colleagues,¹ 94% of Japanese CABG programs would fall to the left of the vertical line (annual volume of 100 procedures) in Figure 1, which would be very low volume by US and Canadian standards.

By comparison, during 2000–2001, the median volume of CABG procedures among STS National Adult Cardiac Database participants was 253,¹³ notably still less than the Leapfrog threshold of 450 procedures. At the high-volume extreme, New York has had a longstanding aggressive approach to monitoring and improving cardiac surgery quality.¹⁶ Between 1997 and 1999, median CABG volume at New York hospitals was 527 procedures (mean, 577 procedures; interquartile range, 331–816 procedures). Only 2.14% of patients undergoing CABG were treated at hospitals performing fewer than 200 such procedures annually,^{10,11} and only about one tenth of New York hospitals had annual CABG volumes of less than 200 procedures.

Given the consistently low volume of most Japanese CABG programs, their overall results will come as a surprise to many. The most complete data source for this study was a survey of 540 programs collected by the Japanese Association for Thoracic Surgery, including almost all programs in the country. The overall mortality rate was 1.9%, and mortality for all volume categories above 41 to 50 procedures per year was less than 2%, which is comparable with rates in most US state and national CABG registries. Mortality rates

TABLE 1. 2004 CABG volumes, numbers of programs, and outcomes for selected states and provinces compared with estimated annual Japanese experience

Region	Population aged ≥ 18 y*	2004 isolated CABG admissions	2004 CABG admissions per 100,000 adults	2004 CABG programs	2004 CABG programs per 1000 CABG admissions	2004 isolated CABG volume, median (min-max)	2004 crude operative [†] mortality (%)	2004 RAMR (%), median (min-max) [‡]
California [§]	26,924,935	19,101	70.9	120	6.3	120 (4–975)	3.29	3.30 (0–12.5)
New Jersey	6,635,222	6177	93.1	17	2.7	290 (102–755)	1.98	NA
Massachusetts [¶]	4,988,309	3986	79.9	14	3.5	287 (101–537)	2.01	2.09 (1.50–3.95)
New York [#]	14,791,841	12,988	87.8	39	3.0	288 (1–1188)	2.09	2.01 (0–3.6)
Ontario ^{**}	9,439,990	7196	76.2	11	1.5	647 (305–945)	1.24	1.11 (0.46–4.11)
Pennsylvania ^{††}	9,635,748	13,359	138.6	60	4.5	195 (56–888)	2.31	NA
Japan (current study) ^{††}	105,943,707	20,000 (††)	18.9	540+	27	28 (0.25–293) (††)	1.92 (††)	NA

CABG, Coronary artery bypass grafting; RAMR, risk-adjusted mortality rate; NA, not applicable.

* For the US States: US Census Bureau Population Estimates 2006 (<http://www.census.gov/popest/states/asrh/SC-EST2006-01.html>); for Ontario: Statistics Canada, Canada's National Statistics Agency, Age and Sex for the Population of Canada, Provinces, Territories, Census Metropolitan Areas and Census Agglomerations, 2001 and 2006 Censuses (<http://www12.statcan.ca/english/census06/data/topics/RetrieveProductTable.cfm?ALEVEL=3&APATH=3&CATNO=&DETAIL=0&DIM=0&DS=99&FL=0&FREE=0&GAL=0&GC=99&GK=NA&GRP=1&IPS=&METH=0&ORDER=1&PID=88984&PTYPE=88971&RL=0&S=1&ShowAll=No&StartRow=1&SUB=0&Temporal=2006&Theme=66&VID=0&VNAMEE=&VNAMEF=&GID=837983>); for Japan: Statistics Bureau, 2005 Population Census, First Basic Complete Tabulation (<http://www.stat.go.jp/English/data/kokusei/2005/kihon1/00/hyodai.htm>).

[†] Operative mortality was defined as follows: for California, Massachusetts, Ontario, and Pennsylvania, all deaths occurring within 30 days of surgical intervention, regardless of where the patient died; for New Jersey, all deaths up to 30 days after surgical intervention or deaths occurring during the hospital stay in which the operation was performed, regardless of the number of days after the procedure; for New York, all deaths within the hospitalization, all discharges (alive or dead) to hospice care except those still alive at 30 days, and all other 30-day deaths.

[‡] New Jersey and Pennsylvania present risk-adjusted mortality rates graphically and do not report specific numbers.

[§] California CABG Outcomes Reporting Program (2007). Coronary Artery Bypass Graft Surgery in California: 2003–2004 Hospital and Surgeon Data. CA Office of Statewide Planning and Development; http://www.oshpd.ca.gov/HID/Products/Clinical_Data/CABG/03-04fullreport.pdf.

^{||} New Jersey Department of Health and Senior Services (May 2007). Cardiac Surgery in New Jersey 2004. Office of the Commissioner; <http://www.nj.gov/health/healthcarequality/documents/cardconsumer04.pdf>.

[¶] Massachusetts Data Analysis Center (October 2006). Adult Coronary Artery Bypass Graft Surgery in the Commonwealth of Massachusetts: January 1, 2004–December 31, 2004. Department of Health Care Policy, Harvard Medical School; <http://www.massdac.org/reports/CABG%202004.pdf>.

[#] New York State Department of Public Health (June 2006). Adult Cardiac Surgery in New York State: 2002–2004; http://www.health.state.ny.us/diseases/cardiovascular/heart_disease/docs/cabg_2002-2004.pdf.

^{**} Guru V, Wang J, Donovan L, Tu JV (June 2006). Report on Coronary Artery Bypass Surgery in Ontario: Fiscal Years 2002–2004; <http://www.ccn.on.ca/memberpdfs/Report-CAB-Surg-Ontario-2002-2004.pdf>.

^{††} Pennsylvania Health Care Cost Containment Council (February 2006). Pennsylvania's Guide to Coronary Artery Bypass Graft Surgery 2004; http://www.phc4.org/reports/cabg/04/docs/cabg_2004report.pdf.

^{††} Miyata and colleagues (current study).¹ 2004 volume estimated from 2001–2004 aggregate data. Number of programs (540) represents, per Miyata and colleagues, almost all CABG programs in Japan. Crude mortality rate is based on 2001–2004 aggregate data. Risk-adjusted results are not presented because they are based on only a small sample of overall programs.

for programs with annual volumes of less than 41 to 50 procedures ranged from 2.42% to 3.15%.

Because these voluntary survey data lacked both adequate risk adjustment and careful audit, the authors also studied a small subset of 36 highly selected Japanese programs that contributed data to the Japanese Adult Cardiovascular Surgery Database, a clinical registry modeled after the STS National Adult Cardiac Database. Table 2 of their article demonstrates that patients in these programs had a distribution of risk factors not dissimilar to what would be observed in many US CABG registries. Unadjusted 30-day mortality was 1.88%, and operative mortality (including in-hospital deaths occurring after 30 days) was 2.55%. Risk-adjusted 30-day mortality was 1.50% for programs with an annual volume of 51 or more procedures and 2.14% for hospital volumes of 31 to 50 procedures.

What Do the Japanese Results Tell Us?

These aggregate data illustrate the feasibility of achieving good overall performance at low volumes, but they do not address the problem of accurately measuring individual hospital performance based on small samples. Although individual hospital volumes and outcomes are not provided by the authors, one would presume that at median volumes of 28 procedures per year, random sampling variation alone could result in mortality rates from zero to greater than 10% for some programs, regardless of their underlying true performance. With volumes and sampling variability in this range, there is virtually no practical way to monitor quality in any meaningful time frame.

What are the implications of these generally favorable results for policymakers in Japan and the United States? In our view, despite the reasonable overall results, this study certainly should not be interpreted as justification for reduced vigilance of low-volume programs. Rather, it is just another example, albeit a dramatic one, of how low-volume programs can often function quite well. This study does not address the issue of identifying individual high- and low-performing, low-volume programs.

The Japanese study also leaves many other important questions unanswered. For example, there are few if any truly high-volume programs in this study with which to compare the performance of their otherwise low-volume centers. Overall Japanese CABG mortality rates seem reasonable, but are they optimal given their demographics and current resources? If Japan had a number of true high-volume programs, would their mortality rates for the same patient population be even lower than those currently being reported?

How Can Low-volume Programs Perform at High Levels?

These seemingly good results should also lead us to inquire more deeply into the structures and processes of care at these

low-volume Japanese programs. Even with the obvious caveats regarding sample size and risk adjustment, there is no compelling evidence in these results or elsewhere in the literature of a CABG mortality crisis in Japan. Perhaps there are lessons from this extreme example that might be applicable to less extreme but smaller programs in other countries. To some extent there is truth in the axiom that practice makes perfect, and some level of repetition is essential for any complex task. But repetition is not the only or perhaps even the best path to high performance. Countless repetitions of suboptimal practices only reinforce those practices and leads to no improvement whatsoever. Identification of optimal practices affords the opportunity for substantial learning with every repetition, even if the overall number of repetitions is smaller.

Why have so many programs been able to perform at least reasonably well in this very low-volume environment? Are there unmeasured differences in case mix or selection criteria compared with the US experience? Is there a proportionately higher volume of other types of cardiac surgery, including CABG combined with other procedures, that maintains both technical proficiency for individual surgeons and also effective team functioning? Are there particular surgical techniques, perfusion methodologies, or standardized perioperative care routines that have enhanced the overall outcomes of their patients?

Team functioning can affect CABG outcomes to a greater extent than in other procedures. Cardiac surgery is a team effort, a complex interaction of surgeons, assistants, nurses, anesthesiologists, and perfusionists. Low-volume surgeons have better results when operating at high-volume hospitals,^{8,11} presumably because they benefit from their standardized processes and team functioning. By the same token, even a higher-volume surgeon might not function optimally when working with an unfamiliar team. In a recently reported California experience, much of the salutary effect of being a higher-volume surgeon was negated if the surgeon did not perform these cases in the same institution (and presumably with the same team).¹⁷

In the Japanese experience most surgeons work with their own teams and function in only one hospital. Many cases are staffed by 2 attending surgeons, and the authors also note the importance of physician oversight of postoperative intensive care unit care. Finally, because of their insurance system, there is less pressure for early discharge, which might reduce the frequency and adverse effect of unrecognized late complications and readmissions.

Recommendations

In light of past research and current findings, we can envision a number of approaches to low-volume CABG surgery that might have applicability both in Japan and elsewhere.

Extreme Low-volume CABG Surgery: Consolidation or Regionalization

Is there a lower acceptable limit for CABG volume? Despite our strong support of direct outcomes measurement and general skepticism of volume thresholds, it is intuitively difficult to believe that optimal CABG results can be achieved in an institution performing a few procedures a month, the situation in many Japanese programs. Perhaps there are unique geographic, demographic, cultural, or political considerations that support the perpetuation of the most extreme low-volume Japanese programs, but these should be individually reviewed. Within this group, it is certainly possible that some programs perform well. However, at such extreme low volumes, there is no possibility of accruing sufficient data to reliably measure performance, at least in a reasonable time frame. Consolidation of the most extreme low-volume programs would be, in our opinion, a significant step toward a more rational CABG delivery system in Japan (and elsewhere). How to define this category of programs is challenging, but in many areas fewer than 100 to 125 isolated CABG procedures annually might be a reasonable starting point for discussion. The exception might be a center performing a large number of other cardiac procedures combined with CABG.

Low-volume CABG Surgery: Improving Performance Assessment

There are many smaller programs above the extreme low end of the volume spectrum, and for them it is essential to develop better, more comprehensive, and timely methods to monitor performance. This will necessitate uniform adoption of some currently available methods, as well as the implementation of some more innovative approaches.

Clinical data registry

Although participation in a clinical data registry is important for all cardiac surgery programs, it should be absolutely mandatory for lower-volume programs to maximize the available information regarding patient case mix, appropriateness of surgical indications, and risk-adjusted performance. In Japan, the Japanese Adult Cardiovascular Surgery Database would seem to be an appropriate instrument with which to implement such a program, particularly given its established mechanisms for audit and validation.

Statistical methodologies

For low-volume programs, performance estimates should be based on multiple years of aggregate data and can be reported as a rolling average. This provides larger sample sizes, albeit at the expense of using some data that are several years old and perhaps less relevant to current conditions. The use of hierarchical statistical models is also recommended to address sample size and clustering issues.¹⁸⁻²⁰

Several graphical methods can also aid in monitoring program performance. Funnel plots have been advocated for performance measurement¹⁴ because they explicitly depict the increasing random statistical uncertainty of a binomial event at small sample sizes. The results from low-volume programs look much less anomalous when viewed from this perspective.

Another graphical approach to monitoring performance, the CUSUM (Cumulative Sum) chart and its variants,²¹ has also been used increasingly in recent years. These methods provide sensitive, real-time monitoring with the potential for earlier detection of deteriorating trends in performance, and they might be less dependent on the accrual of large sample sizes.

Composite measures

Composite measures of CABG quality have recently been developed and implemented by the STS,^{22,23} and these might be particularly advantageous in following small programs that have a correspondingly low number of mortality end points. Because they contain more end points encompassing multiple domains of care (not just mortality), such composites are useful in assessing and comparing quality. Additional end points include both morbidity outcomes measures and process measures, the latter including use of internal thoracic artery grafts and medications proved to reduce long-term cardiovascular risk. The STS composite CABG measure has been shown to enhance the ability to differentiate performance among providers.

Appropriateness

Given the pressure to attain volume thresholds, it is particularly important to monitor procedures at lower-volume programs for appropriateness. This can be done by using standard criteria established by the American College of Cardiology/American Heart Association.

Patient satisfaction

Measures of patient satisfaction are becoming increasingly important, such as the Consumer Assessment of Healthcare Providers and Systems program. Such measures can be particularly useful in assessing the relative value to patients of having surgical intervention in their smaller local hospital versus traveling to a larger tertiary center.

Direct expert review

Finally, in some situations regulators might determine that a particular low-volume program must undergo case-by-case monitoring by an external expert panel as the quid pro quo for continued licensure. This expert committee could meet on a regular basis to review both appropriateness and outcomes for each case.

Low-volume CABG Surgery: Process and Structural Aids to High Performance

Adherence to best practice guidelines

In addition to using more reliable and comprehensive measures of performance, smaller programs must have strong incentives to use established best practices. Based on practices developed at larger institutions, these might help to offset their relatively smaller experience.

Case selection

Except for emergencies in which patient transfer is not feasible, smaller programs should be highly selective in the type of cases they perform. Although the individual surgeons might be skilled and experienced in more complex operations, the experience of the team will likely be limited, as will the other hospital resources necessary to care for such patients perioperatively.

Sponsorship and oversight by tertiary centers

Ideally, low-volume programs should not function in isolation. When feasible, they should be sponsored by larger tertiary centers that share the responsibility for ensuring and improving their quality. Standardized processes of care can be directly imported from the tertiary center, and there might be periodic exchange of staff to bring new ideas and techniques. Teaching conferences and lectures can be scheduled regularly, either live or by means of videoconferencing, and there might even be resident rotations from the tertiary center to the low-volume program. The low-volume center may enjoy some volume purchasing advantages because of its affiliation with the larger center. Sponsoring institutions share joint responsibility with the low-volume center for staff credentialing, scrutiny of outcomes, and remediation when appropriate. Finally, the low-volume center has an established referral pathway for more complex and severely ill patients. Because they are part of the larger program, there is no incentive to retain cases for which they are not equipped (patient-program mismatch). A prototype for such an arrangement has been in place in Massachusetts for a number of years and has functioned quite well.

Team functioning

Small programs should generally be restricted to one surgical practice group to minimize the potentially adverse effect of competition for a limited number of cases. This structure also maximizes the joint experience of the single surgical team, an extremely important consideration, particularly given the recent report from California.¹⁷ It may be useful to have two attending surgeons scrub, particularly for more difficult cases, apparently a common practice in Japan. This maximizes the experience of both surgeons and provides additional peer assistance.

Team functioning and cohesiveness can also be fostered through crew resource management training, simulations, and regular team visits to tertiary centers.

Summary

Volume is only a proxy for outcomes, such as risk-adjusted mortality, and the volume–outcome relationship holds true only on average.⁹ Because of the weak relationship of CABG outcomes to volume and the large sampling variation of observed mortality at low volume, attempts to improve CABG performance primarily through volume thresholds are problematic, except at the extreme. Direct outcomes measurement is a much more reliable approach for both accountability and consumer guidance, and it is more predictive of subsequent program performance.²⁴ From this perspective, the greatest advantage of volume thresholds is to increase sample sizes to enable more precise assessment of risk-adjusted outcomes,²⁵ the real metric on which we should be focusing.

Many low-volume programs perform at a high level. Although it is appropriate to discourage extremely low-volume programs, many lower-volume centers not meeting Leapfrog or similar thresholds do provide high value to the public. Approaches should be considered that optimize the functioning of such units, and more comprehensive techniques should be used to facilitate performance monitoring in lower-volume settings.

References

1. Miyata H, Motomura N, Ueda Y, Matsuda H, Takamoto S. Effect of procedural volume on outcome of CABG surgery in Japan: implication toward public reporting and minimal volume standards. *J Thorac Cardiovasc Surg.* 2008;135:1306-12.
2. Urbach DR, Austin PC. Conventional models overestimate the statistical significance of volume–outcome associations, compared with multilevel models. *J Clin Epidemiol.* 2005;58:391-400.
3. Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. *N Engl J Med.* 1979;301:1364-9.
4. Birkmeyer JD, Siewers AE, Finlayson EVA, Stukel TA, Lucas FL, Batista I, et al. Hospital volume and surgical mortality in the United States. *N Engl J Med.* 2002;346:1128-37.
5. Shahian DM, Normand SL. The volume–outcome relationship: from Luft to Leapfrog. *Ann Thorac Surg.* 2003;75:1048-58.
6. Welke KF, Barnett MJ, Vaughan Sarrazin MS, Rosenthal GE. Limitations of hospital volume as a measure of quality of care for coronary artery bypass graft surgery. *Ann Thorac Surg.* 2005;80:2114-9.
7. Rathore SS, Epstein AJ, Volpp KG, Krumholz HM. Hospital coronary artery bypass graft surgery volume and patient mortality, 1998-2000. *Ann Surg.* 2004;239:110-7.
8. Hannan EL, Kilburn H Jr, Bernard H, O'Donnell JF, Lukacik G, Shields EP. Coronary artery bypass surgery: the relationship between in-hospital mortality rate and surgical volume after controlling for clinical risk factors. *Med Care.* 1991;29:1094-107.
9. Hannan EL. The relation between volume and outcome in health care. *N Engl J Med.* 1999;340:1677-9.
10. Wu C, Hannan EL, Ryan TJ, Bennett E, Culliford AT, Gold JP, et al. Is the impact of hospital and surgeon volumes on the in-hospital mortality rate for coronary artery bypass graft surgery limited to patients at high risk? *Circulation.* 2004;110:784-9.
11. Hannan EL, Wu C, Ryan TJ, Bennett E, Culliford AT, Gold JP, et al. Do hospitals and surgeons with higher coronary artery bypass graft surgery volumes still have lower risk-adjusted mortality rates? *Circulation.* 2003;108:795-801.

12. Nallamothu BK, Saint S, Ramsey SD, Hofer TP, Vijan S, Eagle KA. The role of hospital volume in coronary artery bypass grafting: is more always better? *J Am Coll Cardiol*. 2001;38:1923-30.
13. Peterson ED, Coombs LP, DeLong ER, Haan CK, Ferguson TB. Procedural volume as a marker of quality for CABG surgery. *JAMA*. 2004;291:195-201.
14. Spiegelhalter DJ. Funnel plots for comparing institutional performance. *Stat Med*. 2005;24:1185-202.
15. Dimick JB, Welch HG, Birkmeyer JD. Surgical mortality as an indicator of hospital quality: the problem with small sample size. *JAMA*. 2004;292:847-51.
16. Hannan EL, Kilburn H Jr, Racz M, Shields E, Chassin MR. Improving the outcomes of coronary artery bypass surgery in New York State. *JAMA*. 1994;271:761-6.
17. Carey J, Parker J. The occasional open heart surgeon revisited. Available at: <http://www.aats.org/annualmeeting/Abstracts/2007/37.html>. Accessed November 5, 2007.
18. Normand S-LT, Shahian DM. Statistical and clinical aspects of hospital outcomes profiling. *Stat Sci*. 2007;22:206-26.
19. Shahian DM, Normand SL, Torchiana DF, Lewis SM, Pastore JO, Kuntz RE, et al. Cardiac surgery report cards: comprehensive review and statistical critique. *Ann Thorac Surg*. 2001;72:2155-68.
20. Shahian DM, Torchiana DF, Shemin RJ, Rawn JD, Normand SL. Massachusetts cardiac surgery report card: implications of statistical methodology. *Ann Thorac Surg*. 2005;80:2106-13.
21. Rogers CA, Reeves BC, Caputo M, Ganesh JS, Bonser RS, Angelini GD. Control chart methods for monitoring cardiac surgical performance and their interpretation. *J Thorac Cardiovasc Surg*. 2004;128:811-9.
22. O'Brien SM, Shahian DM, DeLong ER, Normand SL, Edwards FH, Ferraris VA, et al. Quality measurement in adult cardiac surgery: part 2—statistical considerations in composite measure scoring and provider rating. *Ann Thorac Surg*. 2007;83(suppl):S13-26.
23. Shahian DM, Edwards FH, Ferraris VA, Haan CK, Rich JB, Normand SL, et al. Quality measurement in adult cardiac surgery: part 1—conceptual framework and measure selection. *Ann Thorac Surg*. 2007;83(suppl):S3-12.
24. Birkmeyer JD, Dimick JB, Staiger DO. Operative mortality and procedure volume as predictors of subsequent hospital performance. *Ann Surg*. 2006;243:411-7.
25. Luft HS. Better for whom? Policy implications of acting on the relation between volume and outcome in coronary artery bypass grafting. *J Am Coll Cardiol*. 2001;38:1931-3.