

# Stroke

American Stroke  
Association<sup>SM</sup>

JOURNAL OF THE AMERICAN HEART ASSOCIATION

A Division of American  
Heart Association



## **To What Extent Should Quality of Care Decisions Be Based on Health Outcomes Data?: Application to Carotid Endarterectomy**

Gregory Samsa, Eugene Z. Oddone, Ronnie Horner, Jennifer Daley, William Henderson and David B. Matchar

*Stroke* 2002;33:2944-2949; originally published online Oct 10, 2002;

DOI: 10.1161/01.STR.0000038095.20079.F6

Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75214

Copyright © 2002 American Heart Association. All rights reserved. Print ISSN: 0039-2499. Online ISSN: 1524-4628

The online version of this article, along with updated information and services, is located on the World Wide Web at:

<http://stroke.ahajournals.org/cgi/content/full/33/12/2944>

Subscriptions: Information about subscribing to Stroke is online at  
<http://stroke.ahajournals.org/subscriptions/>

Permissions: Permissions & Rights Desk, Lippincott Williams & Wilkins, a division of Wolters Kluwer Health, 351 West Camden Street, Baltimore, MD 21202-2436. Phone: 410-528-4050. Fax: 410-528-8550. E-mail:  
[journalpermissions@lww.com](mailto:journalpermissions@lww.com)

Reprints: Information about reprints can be found online at  
<http://www.lww.com/reprints>

# To What Extent Should Quality of Care Decisions Be Based on Health Outcomes Data?

## Application to Carotid Endarterectomy

Gregory Samsa, PhD; Eugene Z. Oddone, MD; Ronnie Horner, PhD; Jennifer Daley, MD; William Henderson, PhD; David B. Matchar, MD

**Background and Purpose**—Most quality improvement methods implicitly assume that facilities with high complication rates are likely to have substandard processes of care, a stable characteristic that, in the absence of intervention, will persist over time. We assessed the extent to which this holds true for carotid endarterectomy.

**Methods**—Using data from the Department of Veterans Affairs National Surgical Quality Improvement Project, we classified facilities on the basis of 30-day complications of carotid endarterectomy (stroke, myocardial infarction, death) during 1994 to 1995 (period 1, n=3389) and then compared these groups of facilities for complication rates during 1996 to 1997 (period 2, n=4453).

**Results**—Despite wide variation in facility-specific complication rates, the correlation between rates in periods 1 and 2 was low (Spearman correlation coefficient, 0.04;  $P=0.01$ ). Facility-specific rates did not show greater correlation when we examined only facilities with higher volumes patients in different clinical categories (asymptomatic, transient ischemic attack, stroke). Comorbid illness profiles were similar between the 2 time periods.

**Conclusions**—Most of the facility-specific differences in complication rates in period 1 were not maintained into period 2. Many apparent quality improvement problems may not be as large as they first appear, especially when based on few complications per facility. The inability, in practice, to estimate complication rates at a high degree of precision is a fundamental difficulty for clinical policy making regarding procedures with complication rates such as carotid endarterectomy. (*Stroke*. 2002;33:2944-2949.)

**Key Words:** carotid endarterectomy ■ complications ■ quality improvement

For patients with high-grade carotid stenosis, the most important factor in the decision of whether to perform carotid endarterectomy (CEA) is the rate of major surgical complications.<sup>1</sup> This risk depends on patient factors, surgeon factors, hospital- and other system-of-care-related factors, and unexplained factors, the last category often predominating.

Recognizing that sufficiently low complication rates are central to the ultimate effectiveness of CEA, a typical hospital-level quality improvement (QI) initiative might choose to compare surgeons and/or to compare the hospital as a whole against an external benchmark. Because most surgeons perform relatively few procedures, the assessment is usually made at the level of the hospital. Depending on available data, the comparison of the hospital-wide complication rate to the external benchmark might first be made on raw (unadjusted) rates and then followed up through the use

of risk-adjusted rates that take case-mix into account. Because most variation in complication rates is unexplained, the impact of risk adjustment is often small; in this case, adjusted and unadjusted analyses will yield similar conclusions.

When the perspective of a healthcare system is taken, the usual procedure is to calculate complication rates for each facility and then to base decisions concerning QI on these observed rates. Indeed, Luft and colleagues<sup>2</sup> championed this line of investigation in their landmark article documenting the association between surgical volume and operative mortality. One variant on this procedure is to focus on only the “best” (ie, those with the lowest complication rates) and the “worst” (ie, those with the highest complication rates) hospitals. Teams of experts might be sent to both sets of facilities under the assumption that when process of care at the “best” facilities differs from those at the “worst,” changes should be made at the latter based on the former. Another variation is to

Received May 21, 2001; final revision received June 19, 2002; accepted July 3, 2002.

From the Department of Biostatistics and Bioinformatics (G.S.), Center for Clinical Health Policy (G.S., D.B.M.), and Division of General Internal Medicine, Department of Medicine (E.Z.O., R.H., D.B.M.), Duke University Medical Center; Center for Health Services Research in Primary Care, Durham VA Medical Center, Durham, NC (E.Z.O., R.H., D.B.M.); Institute for Health Policy, Massachusetts General Hospital and Partners Healthcare System, Department of Medicine, Harvard Medical School, and Boston VA Medical Center, Boston, Mass (J.D.); and University of Colorado Health Outcomes Program and Denver VA Medical Center, Denver, Colo (W.H.).

Correspondence to Gregory Samsa, PhD, Duke Center for Clinical Health Policy Research, Suite 220, First Union Building, 2200 W Main St, Durham, NC 27705. E-mail samsa001@mc.duke.edu

© 2002 American Heart Association, Inc.

*Stroke* is available at <http://www.strokeaha.org>

DOI: 10.1161/01.STR.0000038095.20079.F6

intervene at all facilities at which the rates are above a certain cut point, which is often based on a published practice guideline or other evidence.

Critical to the above thinking is the assumption that facilities with high complication rates are likely to have a substandard process of care and that this substandard process of care is a stable characteristic that, in the absence of intervention, will persist over time. On the other hand, if observed differences in complication rates are based mostly on statistical variation (ie, “random noise”), then the complication rates in subsequent time periods will tend to regress to the group mean, thus reducing and perhaps even eliminating the need for intervention.<sup>3</sup> Indeed, in this case, the presumed benefits of intervention will be overstated, and classification of facilities as “having a potential quality problem” is, at best, a needless distraction.

With the above in mind, we analyzed data from a prospective cohort of patients undergoing CEA within Department of Veterans Affairs (VA) medical centers and asked this question: To what degree do hospitals with high rates of CEA complications during 1 time period also have high rates of complications in a subsequent time period?

## Methods

### Data Source

This is a secondary analysis of data compiled by the VA National Surgical Quality Improvement Program (NSQIP) regarding patients who underwent CEA. Briefly, the NSQIP is an ongoing, prospective, observational study of outcomes of major surgery performed at 132 VA medical centers. This report focuses on CEA only. For all operations, including CEA, a standard data set was generated. In particular, trained nurse-reviewers collected preoperative data (eg, preoperative assessment, sociodemographic characteristics, clinical characteristics) both directly and through the VA’s computer system. Intraoperative data (eg, CPT-4 codes, operative times, blood transfusions) were collected through the computer system and then verified by the operating surgeon. The reviewer entered the *International Classification of Disease–9-CM* code for the postoperative diagnosis, and followed up the patient for 30 days postoperatively. Hospital-based follow-up included daily rounding, attending conferences, and interviewing surgical house staff and the nurse-epidemiologist regarding possible nosocomial infections and other complications. The reviewer called the patient 30 days after the procedure and interviewed the patient or family member if the patient was unable to communicate, as appropriate.

The NSQIP follows various surgical procedures, and its data collection protocol is not specific to CEA. In particular, the standard data set does not contain some variables that would have been helpful in predicting outcome of CEA, eg, carotid stenosis and information concerning the timing of any prior stroke. Also, there was no requirement for a comprehensive examination by a neurologist either before or after the surgery, implying that those strokes—presumably minor—that would have become known only after such an examination may have been missed.

Additional details pertaining to NSQIP design and methodology are reported elsewhere.<sup>4–9</sup> We included data from 1994 to 1997 and limited the analysis to facilities performing procedures throughout this period. If a patient underwent >1 CEA during this time, we selected the first.

### Outcome

The primary outcome was major complications of CEA within 30 days after surgery, defined to include any or all of stroke, myocardial infarction, and death (regardless of cause). Because complication rates were relatively low, we did not analyze any of the above

components of the overall complication rate separately. As secondary outcomes, we also report the presence of any complication (most commonly pneumonia, prolonged intubation, reintubation in the postoperative period, urinary tract infection, and cardiac arrest), as well as procedure-related return to the operating room.

### Statistical Analysis

The research question asked whether facilities with high complication rates during 1 period also had higher complication rates in a subsequent period. Accordingly, we divided the data set by year, with period 1 consisting of 1994 to 1995 and period 2 consisting of 1996 to 1997. We then formed a data array with 1 record per facility, having for each period the number of patients at risk, the number of patients with complications, and the observed complication rate. The primary predictor variable was observed complication rate during period 1; the primary outcome variable was observed complication rate during period 2.

The main analysis grouped the facilities according to complication rate during period 1 and then used a  $\chi^2$  test to compare the complication rates in period 2. The period 1 groupings were selected as 0%, >0% to 3% (denoted as 0% to 3%), >3% to 5% (denoted as 3% to 5%), >5% to 7% (denoted as 5% to 7%), and >7%; these groupings correspond to various clinically relevant cut points given in the literature (eg, 3%, 5%, and 7%) and allow comparison of facilities with very low rates (0%) with those with very high rates (>7%).

Statistical comparisons were made with a 1-*df* Mantel-Haenszel  $\chi^2$  test (comparing the complication rates in period 2, accounting for the ordinality of the period 1–based facility groupings) and a Spearman correlation coefficient (quantifying the magnitude of this association). As a technical point, it should be noted that our primary interest was in the absolute magnitude of the differences between the groups in period 2, not necessarily in the probability value that compares these rates. In particular, this probability value tests the null hypothesis that there is no difference whatsoever in outcome rates (ie, that all the variation in period 2 is due the effects of “noise” at time 1). Our interest was not in the question of whether none of the variation in complication rates during period 2 is predictable from the rates in period 1 (this being assessed by the above probability value) but how much of the variation in period 2 is predictable (this being assessed by the analyses described above).

## Results

The number of unique patient records available for analysis was 7842: 3389 during 1994 to 1995 and 4453 during 1996 to 1997. Table 1 presents characteristics of the patients. Most were older white men. Patients with previous history of stroke, previous history of transient ischemic attack, and no recorded history of cerebrovascular symptoms were all well represented.

Table 1 also presents aggregate data on CEA complications. Considering the primary outcome of stroke, myocardial infarction, or death, the complication rate decreased from 3.9% during 1994 to 1995 to 3.3% during 1996 to 1997. Secondary complications decreased by  $\approx$ 50%.

Table 2 groups facilities according to complication rate in period 1. Comparing the facilities with 0%, 0% to 3%, 3% to 5%, 5% to 7%, and >7% complication rates during period 1 shows that the complication rates during period 2 were 2.4%, 2.8%, 3.5%, 3.9%, and 4.5%, respectively. Facilities grouped according to complication rates in period 1 displayed much less variation in complication rates in period 2. The complication rates in period 2 showed a statistically significant ordinal trend ( $P=0.01$ ), but the magnitude of this trend was modest (Spearman correlation coefficient, 0.04).

**TABLE 1. Demographic Characteristics, Clinical Characteristics, and Complications From CEA**

	1994–1995 (n=3389)	1996–1997 (n=4453)
Age (mean), y	67.3 (7.7)	67.7 (8.1)
Male, %	98.4	98.5
Race, %		
White	92.4	90.4
Black	4.8	5.5
Other	2.8	4.0
History, %		
None	36.0	39.3
TIA	26.1	31.3
Stroke	37.9	29.3
Diabetes, %	15.3	17.9
Renal failure, %	0.2	0.8
COPD, %	17.0	16.7
CHF, %	1.5	2.5
Primary outcome, n (%)		
Stroke, MI, death	136 (3.92)	148 (3.28)
Death	43 (1.27)	46 (1.03)
Stroke	70 (2.07)	89 (2.00)
MI	39 (1.15)	33 (0.74)
Secondary outcome, n (%)		
Any	288 (8.50)	211 (4.74)
Related return to operating room	144 (4.25)	170 (3.82)

TIA indicates transient ischemic attack; COPD, chronic obstructive pulmonary disease; CHF, congestive heart failure; and MI, myocardial infarction.

We then examined whether the above finding could plausibly be attributed to facilities with a small volume of patients. Table 2 repeats the analysis after exclusion of the 31 facilities with <20 procedures during period 1. The relationship between complication rates in periods 1 and 2 was no stronger than before ( $P>0.05$ ; Spearman correlation, 0.02).

We then examined whether the above finding could plausibly be attributed to symptom status. Table 2 repeats the analysis after disaggregating the data according to history of cerebrovascular symptoms. In all 3 clinical symptom categories, the relationship between complication rates in periods 1 and 2 was no stronger than before ( $P>0.05$ ; Spearman correlation, 0.02 to 0.05).

Table 3 reports information about case-mix during the 2 study periods. When facilities were separated into categories based on complication rate in period 1, significant differences in case-mix characteristics were apparent; eg, facilities with high complication rates during period 1 had fewer white patients and more patients with diabetes. These differences in case-mix, however, remained similar across the 2 time periods.

## Discussion

We observed a noteworthy amount of variation in complication rates. For example, although  $\approx 12\%$  of facilities reported period 1 complication rates  $>7\%$ , 31 of the 94 facilities had no complications at all. Although there was some tendency

**TABLE 2. CEA Complication Rates Grouped According to Facility Rates in Period 1**

	1994–1995 Group	1994–1995 Rate, % (n)	1996–1997 Rate, % (n)
Overall*			
0%		0 (0/537)	2.4 (25/1027)
0–3%		2.2 (22/1016)	2.8 (31/1101)
3–5%		4.0 (34/844)	3.5 (34/972)
5–7%		6.2 (36/578)	3.9 (33/843)
$>7\%$		9.9 (41/373)	4.5 (23/510)
Excluding facilities with <20 procedures during 1994–1995†			
0%		0 (0/335)	3.1 (13/425)
0–3%		2.2 (22/1016)	2.8 (31/1101)
3–5%		4.0 (34/844)	3.5 (34/972)
5–7%		6.3 (32/510)	4.0 (26/653)
$>7\%$		9.4 (33/352)	4.0 (15/377)
Patients with no history of cerebrovascular symptoms‡			
0%		0 (0/193)	1.7 (8/470)
0–3%		1.0 (4/401)	2.1 (11/535)
3–5%		2.7 (8/297)	1.7 (6/353)
5–7%		3.3 (7/213)	2.5 (9/360)
$>7\%$		9.5 (15/157)	3.6 (9/252)
Patients with history of TIA§			
0%		0 (0/140)	3.2 (8/250)
0–3%		2.6 (7/269)	2.0 (5/244)
3–5%		4.2 (11/262)	5.8 (13/224)
5–7%		6.4 (12/186)	3.7 (7/189)
$>7\%$		10.1 (13/129)	2.9 (3/102)
Patients with history of stroke			
0%		0 (0/203)	2.9 (9/306)
0–3%		3.2 (11/346)	4.7 (15/322)
3–5%		5.6 (15/270)	3.6 (10/282)
5–7%		9.1 (16/176)	5.6 (14/249)
$>7\%$		10.4 (13/125)	6.8 (10/148)

\* $P=0.010$ , Mantel-Haenszel  $\chi^2$ ; Spearman correlation, 0.04.

† $P=0.170$ , Mantel-Haenszel  $\chi^2$ ; Spearman correlation, 0.02.

‡ $P=0.13$ , Mantel-Haenszel  $\chi^2$ ; Spearman correlation, 0.03.

§ $P=0.56$ , Mantel-Haenszel  $\chi^2$ ; Spearman correlation, 0.02.

|| $P=0.062$ , Mantel-Haenszel  $\chi^2$ ; Spearman correlation, 0.05.

for facilities with high complication rates in period 1 to report relatively high complication rates in period 2 (Spearman correlation, 0.04), most of the differences in complication rates in period 1 were not maintained over time.

Several previous studies have documented important variations in complication rates for CEA, including variation by facility characteristics such as the volume of procedures.<sup>10–12</sup> Implicit in the interpretation of facility-to-facility variation is the assumption that complication rates are reasonably stable within individual facilities. Our findings call into question this assumption of stability.

This basic result has a number of possible explanations. First, noting that our analysis uses unadjusted complication rates, these differences might be due to case-mix. However, the conclusions of a case-mix-adjusted analysis will be similar as long as case-mix is not strongly predictive of

**TABLE 3. Demographic and Clinical Characteristics by Facility Complication Rate**

	0%	0–3%	3–5%	5–7%	>7%	<i>P</i>
<b>1994–1995</b>						
Sample size, n	537	1016	844	578	414	
Age (mean), y	66.9 (7.9)	67.6 (7.7)	67.1 (7.5)	67.6 (7.7)	67.0 (7.9)	0.309
Male, %	98.3	98.1	98.7	98.6	98.1	0.845
Race						
White	90.7	94.3	96.0	93.8	80.3	<0.001
Black	7.7	4.2	2.1	4.9	8.1	
Other	1.5	1.5	1.9	1.2	11.5	
History						
None	36.0	39.5	35.8	37.0	38.2	0.035
TIA	26.1	26.5	31.6	32.3	31.4	
Stroke	37.9	34.1	32.6	30.6	30.4	
Diabetes	15.3	12.7	12.3	13.1	19.3	0.006
Renal failure	0.2	0.9	0.8	1.2	0.7	0.416
COPD	17.0	15.3	11.6	22.9	16.7	0.231
CHF	1.5	4.0	1.9	2.2	3.1	0.012
<b>1996–1997</b>						
Sample size, n	1027	1101	972	843	510	
Age (mean), y	67.4 (8.1)	67.4 (8.2)	67.7 (8.0)	68.3 (8.2)	68.2 (8.1)	0.042
Male, %	98.2	98.5	98.8	98.3	98.8	0.840
Race						
White	90.9	93.6	92.0	89.3	81.6	<0.001
Black	7.5	3.9	3.4	5.9	8.6	
Other	1.6	2.5	4.5	4.9	9.8	
History						
None	45.8	48.6	41.1	45.1	50.2	0.036
TIA	24.4	22.2	26.1	23.7	20.3	
Stroke	29.8	29.2	32.8	31.2	29.5	
Diabetes	25.7	25.8	20.0	26.3	24.9	0.007
Renal failure	0.7	0.4	0.2	0.2	1.0	0.182
COPD	16.9	16.0	16.1	19.2	12.5	0.029
CHF	1.7	3.8	1.1	2.2	4.3	<0.001

Abbreviations as in Table 1.

outcome and/or case-mix remains similar over time. A previous analysis of this data set found no statistically significant relationship between case-mix and outcome.<sup>9</sup> Also, the pattern of case-mix remained relatively stable across facilities. Thus, whatever effect case-mix might be having during period 1 was occurring, in roughly similar measure, in period 2. Moreover, the patterns within clinical state (asymptomatic, history of transient ischemic attack, history of stroke) were similar to those overall, providing further confidence that case-mix is an unlikely explanation.

A second possible explanation is that personnel in the facilities that performed poorly during period 1 noticed this and implemented various interventions to improve the quality of care. Although such an explanation cannot be ruled out with respect to the facilities that were initially poor performers, this does not explain why the complication rates increased among those facilities with initially low rates.

A third possible explanation is that low-volume facilities are particularly unstable from a statistical perspective and are

providing much of the “noise” that is being observed in the complication rates. However, a reanalysis of the data excluding those facilities with low volume, ie, <20 procedures during period 1, did not support such an explanation. Here, it should be noted that our comments are not intended as an addition to the now-considerable literature on the volume-outcome relationship in CEA. Compared with this literature, all the facilities considered here have low to moderate volumes, and our analysis does not consider the experience of the facility with similar procedures or the experience of the surgeons in affiliated non-VA facilities. In any event, our results regarding volume and outcome for CEA performed within the VA are consistent with previous reports from the NSQIP.<sup>9</sup>

Finally, we note that although quality improvement initiatives are often taken at the level of the facility, our assessment of group-level complication rates used as its unit of analysis not the facility but instead groups of facilities with similar



complication rates. While intentionally suppressing the individual facility as a unit of aggregation, this approach does at least serve to avoid various statistical problems involved with differential precision for large and small facilities, facilities having no complications, etc, and is consistent with previous analyses of including facilities with patient volumes that are relatively low and variable.<sup>13</sup> However, the Spearman correlations do use the facility as their unit of analysis.

With the above in mind, we believe that the most likely explanation of the results is regression to the mean—ie, the tendency for groups observed to be extreme on a characteristic to become less extreme on remeasurement. This tendency becomes particularly pronounced when the measure in question is “noisy,” ie, has a high degree of intrinsic variability. Such is the case for CEA complication rates because sample sizes for individual facilities are small and the standard error of a complication rate increases with decreasing sample size and because absolute numbers of complications are small, thus magnifying the impact of any particular patient having a complication.

In interpreting these data, it is important to take into account both the strengths and weaknesses of the NSQIP data. The primary strength, compared with a typical observational study, resides in the great attention paid to comprehensiveness and consistency in data collection. The primary weakness is that the data collection protocol was not specific to CEA. Apart from the desirability of including CEA-specific information about such factors as degree of stenosis in the standard NSQIP data set, a data collection protocol focusing on CEA would likely have included comprehensive examination of all patients by neurologists both before and after surgery. The main implication of the data collection protocol actually used is the possible overlooking of some number of strokes, most likely minor, that would have become apparent only on a comprehensive examination. This suggests that the absolute magnitude of the complication rates reported here may not be comparable with studies such as the randomized trials of CEA that included such examinations. Nevertheless, we believe these results to be internally consistent, in the sense that the same data collection procedures were applied in both time periods; thus, conclusions about the general patterns of complication rates within and across facilities should be unaffected.

As mentioned, the population-based NSQIP study is perhaps most noteworthy because the prospective nature of the data collection supports a much more comprehensive assessment of both patient factors and outcomes than is the case for a typical observational study. Within this NSQIP data set, we found that what initially appeared to be dramatic differences between facilities were most likely due, at least in large part, to artifacts induced by statistical variation. From the perspective of continuous QI, our findings are a reminder that, before attempts to intervene to improve a system are made, it is crucial that the level of common-cause variation be understood.<sup>14</sup> Otherwise, it is seductively easy to conclude that what is in fact common-cause variation is due to special causes such as a substandard process of care. Thus, attempts at intervention might not always be necessary, and resources can sometimes be better spent by simply continuing to

observe and better understand the situation before action is taken.

Our findings primarily illustrate 2 points. First, they serve as a caution to policy makers that apparent QI problems may not always be as large as they first appear, especially those involving facilities having low to moderate patient volumes and/or procedures with relatively low complication rates. Second, they serve to illustrate a fundamental dilemma in clinical policy making regarding CEA, namely, that while the decision whether to perform a CEA depends in large part on small absolute differences between complication rates, in practice it is seldom the case that these complication rates can be estimated to the degree of precision required. For example, even though guidelines and decision analyses make a crucial distinction between rates such as 3%, 4%, 5%, 6%, etc, it is often the case that confidence intervals calculated from observed data from a facility will be consistent with each of the above rates and thus will not be sufficient to make such a differentiation. This dilemma is not limited to CEA but also applies to other procedures having low complication rates (and whose appropriateness or lack thereof depends on relatively small absolute differences in these rates).

Like many others, the perspective of quality improvement can be very helpful when users are cognizant of its limitations yet is considerably less effective when extended beyond its natural sphere of application. For CEA, the level of random variation induced by small absolute numbers of complications circumscribes this sphere of application. Recognizing these limitations, it might be noted that in this respect QI is perhaps not so different from the practice of medicine: The options are to intervene immediately or to simply watch and wait, and the challenge is to select the course of action that leads to the most good and the least harm.

## Acknowledgments

Data were provided by the National Surgical Quality Improvement Program, and funding was provided by the Quality Enhancement Research Initiative, both from the Department of Veterans Affairs.

## References

1. Matchar DB, Pauker SG. Endarterectomy in carotid artery disease: a decision analysis. *JAMA*. 1987;258:793–798.
2. Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. *N Engl J Med*. 1979;301:1364–1369.
3. Davis CEA. Regression to the mean. In: Kotz S, Johnson NL, Read CB, eds. *Encyclopedia of Statistical Sciences*. New York, NY: Wiley and Sons; 1986;7:706–708.
4. Daley J, Forbes MG, Young GJ, Charns MP, Gibbs JO, Hur K, Henderson W, Khuri S, for participants in the National VA Surgical Risk Study. Validating risk-adjusted surgical outcomes: site visit assessment of process and structure. *J Am Coll Surg*. 1997;185:341–351.
5. Daley J, Khuri SF, Henderson W, Hur K, Gibbs JO, Barbour G, Demakis J, Irvin G, Stremple JF, Grover F, McDonald G, Passaro E, Hammermeister K, Aust JB, Oprian C, for participants in the National VA Surgical Risk Study. Risk adjustment of the postoperative morbidity rate for the comparative assessment of the quality of surgical care: results of the National Veterans Affairs Surgical Risk Study. *J Am Coll Surg*. 1997;185:328–340.
6. Khuri SF, Daley J, Henderson W, Barbour G, Lowry P, Irwin G, Gibbs J, Grover F, Hammermeister K, Stremple JF, Aust JB, Demakis J, Deykin D, McDonald G, for participants in the National Veterans Administration Surgical Risk Study. The National Veterans Administration Surgical Risk Study: risk adjustment for the comparative assessment of the quality of surgical care. *J Am Coll Surg*. 1995;180:519–531.

7. Khuri SF, Daley J, Henderson W, Hur K, Gibbs JO, Barbour G, Demakis J, Irwin G, Strempel JF, Grover F, McDonald G, Passaro E, Fabri PJ, Spencer J, Hammermeister K, Aust JB, for participants in the National VA Surgical Risk Study. Risk adjustment of the postoperative mortality rate for the comparative assessment of the quality of surgical care: results of the National Veterans Affairs Surgical Risk Study. *J Am Coll Surg.* 1997;185:315–327.
8. Khuri SF, Daley J, Henderson W, Hur K, Demakis J, Aust JB, Chong V, Fabri PJ, Gibbs JO, Grover F, Hammermeister K, Irwin G, McDonald G, Passaro E, Phillips L, Scamman F, Spencer J, Strempel JF, for participants in the National VA Surgical Quality Improvement Program. The Department of Veterans Affairs' NSQIP: the first national, validated, outcome-based, risk-adjusted, and peer-controlled program for the measurement and enhancement of the quality of surgical care. *Ann Surg.* 1998;228:491–507.
9. Khuri SF, Daley J, Henderson W, Hur K, Hossain M, Soybel D, Kizer KW, Aust JB, Bell RH, Chong V, Demakis J, Fabri PJ, Gibbs JO, Grover F, Hammermeister K, McDonald G, Passaro E, Phillips L, Spencer J, Strempel JF, for participants in the VA National Surgical Quality Improvement Program. The relationship of surgical volume to outcome in eight common operations. *Ann Surg.* 1999;230:414–432.
10. Cebul RD, Snow RJ, Pine R, Hertzer NR, Norris DG. Indications, outcomes, and provider volumes for carotid endarterectomy. *JAMA.* 1998;279:1282–1287.
11. Wennberg DE, Lucas FL, Birkmeyer JD, Bredenberg CEA, Fisher ES. Variation in carotid endarterectomy mortality in the Medicare population: trial hospitals, volume, and patient characteristics. *JAMA.* 1998;279:1278–1281.
12. Dudley RA, Johansen KL, Brand R, Rennie DJ, Milstein A. Selective referral to high-volume hospitals: estimating potentially avoidable deaths. *JAMA.* 2000;283:1159–1166.
13. Mitchell JB, Ballard DJ, Whisnant JP, Ammering CJ, Matchar DB, Samsa GP. Using physician claims to identify postoperative complications of carotid endarterectomy. *Health Serv Res.* 1996;31:141–152.
14. Deming WE. *Out of the Crisis*. Cambridge, Mass: Massachusetts Institute of Technology Center for Advanced Engineering Study; 1982.