

Special Article

HOSPITAL VOLUME AND SURGICAL MORTALITY IN THE UNITED STATES

JOHN D. BIRKMEYER, M.D., ANDREA E. SIEWERS, M.P.H., EMILY V.A. FINLAYSON, M.D., THERESE A. STUKEL, PH.D., F. LEE LUCAS, PH.D., IDA BATISTA, B.A., H. GILBERT WELCH, M.D., M.P.H., AND DAVID E. WENNBURG, M.D., M.P.H.

ABSTRACT

Background Although numerous studies suggest that there is an inverse relation between hospital volume of surgical procedures and surgical mortality, the relative importance of hospital volume in various surgical procedures is disputed.

Methods Using information from the national Medicare claims data base and the Nationwide Inpatient Sample, we examined the mortality associated with six different types of cardiovascular procedures and eight types of major cancer resections between 1994 and 1999 (total number of procedures, 2.5 million). Regression techniques were used to describe relations between hospital volume (total number of procedures performed per year) and mortality (in-hospital or within 30 days), with adjustment for characteristics of the patients.

Results Mortality decreased as volume increased for all 14 types of procedures, but the relative importance of volume varied markedly according to the type of procedure. Absolute differences in adjusted mortality rates between very-low-volume hospitals and very-high-volume hospitals ranged from over 12 percent (for pancreatic resection, 16.3 percent vs. 3.8 percent) to only 0.2 percent (for carotid endarterectomy, 1.7 percent vs. 1.5 percent). The absolute differences in adjusted mortality rates between very-low-volume hospitals and very-high-volume hospitals were greater than 5 percent for esophagectomy and pneumonectomy, 2 to 5 percent for gastrectomy, cystectomy, repair of a nonruptured abdominal aneurysm, and replacement of an aortic or mitral valve, and less than 2 percent for coronary-artery bypass grafting, lower-extremity bypass, colectomy, lobectomy, and nephrectomy.

Conclusions In the absence of other information about the quality of surgery at the hospitals near them, Medicare patients undergoing selected cardiovascular or cancer procedures can significantly reduce their risk of operative death by selecting a high-volume hospital. (N Engl J Med 2002;346:1128-37.)

Copyright © 2002 Massachusetts Medical Society.

OVER the past three decades, numerous studies have described higher rates of operative mortality with selected surgical procedures at hospitals where few such procedures are performed (low-volume hospitals).¹⁻⁴ Several recent reviews suggest that thousands of preventable surgical deaths occur each year in the United States because elective but high-risk surgery is performed in hospitals that have inadequate experience with the surgical procedures involved.⁵⁻⁷ As part of a broader initiative aimed at improving hospital safety, a large coalition of private and public purchasers of health insurance — the Leapfrog Group — is encouraging patients undergoing one of five high-risk procedures to seek care at high-volume hospitals.⁸ In the lay media, there has been an emphasis on the importance of experience with particular procedures,^{9,10} and several consumer-oriented Web sites (e.g., <http://www.healthscope.org>) have begun providing patients with information about volume at hospitals near them.

Despite the recent interest in surgical volume, many question the applicability of previous research on volume and outcome to current practice.^{11,12} First, many studies of volume and outcome are outdated. Given that the surgical mortality associated with many procedures has fallen considerably since these studies were conducted,^{13,14} the relative importance of the volume of procedures performed may be declining. Second, most published studies on volume and outcome have used state-level data bases or regional populations that are served by a small number of high-volume centers.⁶ Whether their results are broadly generalizable is uncertain. And finally, although some procedures (e.g., cardiac surgery) have been studied extensively, the relative importance of hospital volume to mortality with

From the Veterans Affairs Outcomes Group, Department of Veterans Affairs Medical Center, White River Junction, Vt. (J.D.B., E.V.A.F., H.G.W.); the Department of Surgery, Dartmouth-Hitchcock Medical Center, Lebanon, N.H. (J.D.B.); the Center for the Evaluative Clinical Sciences, Dartmouth Medical School, Hanover, N.H. (J.D.B., T.A.S., H.G.W., D.E.W.); the Center for Outcomes Research and Evaluation, Maine Medical Center, Portland (A.E.S., F.L.L., I.B., D.E.W.); and the Department of Surgery, University of California, San Francisco (E.V.A.F.). Address reprint requests to Dr. Birkmeyer at the Veterans Affairs Outcomes Group (111B), Veterans Affairs Medical Center, White River Junction, VT 05009, or at john.birkmeyer@dartmouth.edu.

many other high-risk procedures either has not been explored or has been studied in samples that were too small to permit assessment of performance at all meaningful levels of hospital volume.

To address many of these limitations, we studied surgical mortality in the Medicare population, which accounts for the majority of all patients in the United States who undergo high-risk surgery and an even larger majority of those who die after surgery.¹⁵ Using current national data (from 1994 through 1999), we studied the importance of hospital volume to the operative mortality associated with six types of cardiovascular procedures and eight types of major cancer resections.

METHODS

Subjects and Data Bases

We obtained the Medicare Provider Analysis and Review (MEDPAR) files and the denominator files from the Center for Medicare and Medicaid Services for the years 1994 through 1999. These files contain hospital-discharge abstracts for the acute care hospitalizations of all Medicare recipients covered by the hospital care program (Part A). Only patients covered by fee-for-service arrangements are included in the MEDPAR file; thus, our sample excludes the approximately 10 percent of Medicare patients who were enrolled in risk-bearing health maintenance organizations during this period. We excluded patients who were under 65 years of age or over 99 years of age. The study protocol was approved by the institutional review board of Dartmouth Medical School.

Hospital Volume

Patients undergoing each of the 14 procedures examined in our analysis were identified with the use of appropriate procedure codes from the *International Classification of Diseases, 9th Revision, Clinical Modification* (ICD-9-CM).¹⁶ These procedures were selected because they are relatively complex, are associated with a nontrivial risk of operative mortality, and are most often performed on an elective basis.

We focused on the total number of each type of procedure performed at a given hospital (hospital volume), not the total number of procedures involving Medicare recipients (Medicare volume), in order to place our results in the context of the volume standards suggested by the Leapfrog Group⁸ and others. To estimate total volumes, we examined data from the all-payer 1997 Nationwide Inpatient Sample. We determined the proportion of all patients undergoing each procedure who were covered by Medicare; the proportion ranged from 43 percent (for nephrectomy) to 75 percent (for carotid endarterectomy). To estimate the total volume at individual hospitals, we divided the observed Medicare volume (the total number of each type of procedure performed on Medicare patients during the six-year study period) by these procedure-specific proportions.

Hospital volume, expressed as the average number of procedures per year, was first evaluated as a continuous variable. To simplify the presentation of our results, however, we also created categorical variables, defining five categories of hospital volume: very low, low, medium, high, and very high. For each procedure, the hospitals were ranked in order of increasing total hospital volume, and then five volume groups were defined by the selection of whole-number cutoff points for annual volume that most closely sorted the patients into five groups of equal size (quintiles). The cutoff points were established before mortality was examined in order to avoid selecting cutoff points that could maximize the associations between volume and

outcome.¹⁷ To reflect most accurately the overall institutional experience with each type of operation, we combined the replacement of aortic and mitral valves (into the single category of heart-valve replacement) and lobectomy and pneumonectomy (into the category of lung resection) in determining hospital volume. However, the outcomes of these procedures were assessed separately.

Assessment of Outcomes

In creating cohorts for the analysis of outcomes, we applied several restrictions in order to increase the homogeneity of the study samples and thus minimize the potential for confounding by case mix. For the eight types of major cancer resections, we excluded patients without an accompanying cancer-diagnosis code (related to the index procedure). Patients undergoing repair of an abdominal aortic aneurysm were excluded if they had a diagnosis or procedure code suggesting rupture of the aneurysm, thoracoabdominal aneurysm, or both. Patients undergoing coronary-artery bypass grafting were excluded if they simultaneously underwent valve replacement.

Our primary outcome measure was operative mortality, defined as the rate of death before hospital discharge or within 30 days after the index procedure. Because a large proportion of surgical deaths before discharge occurred more than 30 days after surgery, we decided that 30-day mortality alone would not adequately reflect true operative mortality. Because the length of stay did not vary systematically according to hospital volume, the inclusion of late, in-hospital deaths would not be expected to bias our results. Moreover, associations between volume and outcome were largely unchanged when we repeated our analyses using 30-day mortality alone.

Statistical Analysis

We used multiple logistic regression to examine relations between hospital volume and operative mortality, with adjustment for characteristics of the patients.¹⁸ We used the patient as the unit of analysis, with volume measured at the hospital level. We first fitted separate models for each procedure against the logarithm of volume to establish the general form of the relation. We then fitted models against the quintiles of volume for each procedure.

We adjusted for age group (65 to 69 years, 70 to 74 years, 75 to 79 years, 80 to 84 years, or 85 to 99 years), sex, race (black or non-black), and their interactions, as well as the year of the procedure, the relative urgency of the index admission (elective, urgent, or emergency), the presence of coexisting conditions, and mean income from Social Security.¹⁸ This last measure was assessed at the ZIP Code level (on the basis of the 1990 Census file) because patient-level information on socioeconomic status is not available.

Coexisting conditions were identified with the use of information from the index admission and any other admissions that had occurred within the preceding six months. Relative to low-volume hospitals, high-volume hospitals treat a larger number of patients who have been transferred or referred from other centers. To minimize the possibility of bias due to the identification of more previous admissions (and thus more coexisting conditions) at high-volume centers, we excluded information on coexisting conditions identified at previous admissions that occurred within two weeks before the index hospitalization. For the purposes of risk adjustment, coexisting conditions (identified by their appropriate ICD-9-CM codes) were compiled into a Charlson score (the number of coexisting conditions, weighted according to their relative effects on mortality),^{19,20} which was modified to exclude conditions that were likely to reflect either the primary indication for surgery or postoperative complications.^{21,22} We also explored two alternative approaches to incorporating coexisting conditions into our risk-adjustment models: including Charlson scores with weights derived empirically for each procedure and including coexisting conditions individually by inserting into each model each condition that was present in at least 2 percent of the patients. Because all three approaches yielded virtually identical results, we report only those from the model that used the Charlson score with published weights.¹⁹

We used overdispersed binary logistic models to adjust for clustering of deaths within hospitals.²³ The net effect was to increase the width of the confidence intervals between 2 percent (cystectomy) and 44 percent (lobectomy), with a mean increase of 25 percent. We computed adjusted mortality rates on the basis of the average values of the characteristics of the patients by back-transforming predicted mortality from the logistic model. Our final risk-adjustment models had intermediate discriminative ability, with C statistics ranging from 0.60 (for pneumonectomy) to 0.71 (for nephrectomy). All P values are two-tailed.

Because the Medicare files used for this analysis reflect the use of procedures among patients with fee-for-service arrangements for health care, we may have underestimated hospital volume in regions of the country that had a high penetration of Medicare managed care during the study period (mainly southern California and the Southwest). For this reason, we repeated our analyses after restricting our data set to hospital-referral regions with a penetration of Medicare managed care of less than 10 percent. Because the adjusted odds ratios for death associated with hospital volume changed negligibly as a result of this restriction, these data are not presented.

RESULTS

Between 1994 and 1999, approximately 2.5 million Medicare patients underwent 1 of the 14 cardiovascular or cancer-related procedures that we studied. The criteria used to define the five strata of hospital volume varied markedly according to procedure, reflecting the relative frequency with which each is performed (Table 1). Medicare volume and total volume for the 14 procedures were highly correlated at the hospital level (overall correlation coefficient, 0.97).

The age and sex of the patients did not vary consistently among strata of hospital volume (Table 2). However, for most of the 14 procedures, black patients were more likely to undergo surgery at a lower-volume hospital. For most procedures, Charlson scores tended to be slightly higher at higher-volume hospitals. However, patients were more likely to be admitted non-electively at lower-volume hospitals. This trend was more apparent with respect to several cancer-related resections (e.g., esophagectomy) than with respect to cardiovascular procedures.

When it was assessed as a continuous (logarithmic) variable, hospital volume was related to both observed and adjusted operative mortality rates for all 14 procedures ($P < 0.001$ for all trends). In terms of odds ratios for death, adjustment for characteristics of the patients attenuated the associations between volume and outcome moderately for carotid endarterectomy, colectomy, gastrectomy, esophagectomy, and pulmonary lobectomy (Table 3). Risk adjustment had negligible effect with respect to the other procedures.

In terms of absolute differences in adjusted mortality rates, the importance of hospital volume varied markedly according to the type of procedure (Fig. 1 and 2). For example, for pancreatic resection, adjusted mortality rates at very-low-volume hospitals were 12.5 percent higher than at very-high-volume hospitals (16.3 percent vs. 3.8 percent) (Fig. 2A). Relatively large differences in risk were also observed for esoph-

agectomy (11.9 percent) and pneumonectomy (5.4 percent). Absolute differences in adjusted mortality rates between very-low-volume and very-high-volume hospitals were between 2 percent and 5 percent for gastrectomy, cystectomy, repair of a nonruptured abdominal aortic aneurysm, and aortic- and mitral-valve replacement, and the differences were less than 2 percent for coronary-artery bypass grafting, lower-extremity bypass, colectomy, lobectomy, and nephrectomy. The absolute difference in mortality between very-low-volume and very-high-volume hospitals was smallest for carotid endarterectomy (1.7 percent vs. 1.5 percent).

Relations between volume and outcome in the intermediate strata of hospital volume also varied widely according to the type of procedure (Fig. 1 and 2). For several types of procedure (including coronary-artery bypass grafting, valve replacement, and pancreatic resection), mortality declined monotonically with each stratum of increasing hospital volume. For others (including elective repair of an abdominal aortic aneurysm, gastrectomy, and pneumonectomy), differences in mortality were most apparent at the extremes of volume, whereas hospitals in the intermediate-volume strata had similar mortality rates.

DISCUSSION

In this large, national study, higher-volume hospitals had lower operative mortality rates for six types of cardiovascular procedures and eight types of major cancer resections. However, the absolute magnitude of the relation between volume and outcome varied markedly among the types of procedures. Dramatic differences in mortality between very-low-volume and very-high-volume hospitals were observed for pancreatic resection and esophagectomy (more than 12 percent, in absolute terms), whereas relatively small differences in mortality (1 percent or less) were found for 3 of the 14 procedures examined in our analysis. These findings suggest the relative importance of hospital volume for individual patients who are considering where to undergo various procedures. From the public health perspective, however, one must also consider the total number of patients who undergo each procedure. For example, in the case of coronary-artery bypass grafting (for which volume had a moderate effect but which is very common), 314 deaths would be averted in the United States each year if the mortality rate at very-low-volume hospitals were reduced to the rate at very-high-volume centers. Conversely, in the case of pancreatic resection (for which volume had a very large effect but which is performed infrequently), lowering the mortality rate at very-low-volume centers to that observed at very-high-volume centers would avert only 32 deaths annually.

We believe that our results reflect real differences

HOSPITAL VOLUME AND SURGICAL MORTALITY

TABLE 1. DISTRIBUTION OF PATIENTS AND HOSPITALS AMONG QUINTILES OF VOLUME FOR THE 14 PROCEDURES.*

PROCEDURE	HOSPITAL VOLUME				
	VERY LOW	LOW	MEDIUM	HIGH	VERY HIGH
Coronary-artery bypass grafting					
No. of Medicare patients	188,176	157,060	197,322	181,427	177,682
No. of hospitals	595	181	149	91	52
Average no. of procedures/yr (57% Medicare)	<230	230-348	349-549	550-849	>849
Heart-valve replacement					
Aortic					
No. of Medicare patients	30,377	30,329	31,590	29,607	29,707
No. of hospitals	639	198	123	70	39
Mitral					
No. of Medicare patients	12,526	13,010	13,560	12,603	13,236
No. of hospitals	620	198	123	70	39
Average no. of procedures/yr (62% Medicare)	<43	43-74	75-119	120-199	>199
Carotid endarterectomy					
No. of Medicare patients	101,319	85,009	102,038	95,587	95,336
No. of hospitals	2,013	404	291	180	102
Average no. of procedures/yr (75% Medicare)	<40	40-69	70-109	110-164	>164
Lower-extremity bypass					
No. of Medicare patients	53,974	53,080	53,221	52,958	50,347
No. of hospitals	2,068	502	305	204	105
Average no. of procedures/yr (70% Medicare)	<22	22-39	40-60	61-94	>94
Elective repair of abdominal aortic aneurysm					
No. of Medicare patients	27,970	27,273	29,029	28,884	27,421
No. of hospitals	1,900	426	257	156	80
Average no. of procedures/yr (70% Medicare)	<17	17-30	31-49	50-79	>79
Colectomy					
No. of Medicare patients	63,386	58,720	62,789	59,398	59,992
No. of hospitals	2,856	727	486	319	199
Average no. of procedures/yr (59% Medicare)	<33	33-56	57-84	85-124	>124
Gastrectomy					
No. of Medicare patients	6,324	6,287	6,871	6,194	6,268
No. of hospitals	2,010	636	419	239	119
Average no. of procedures/yr (54% Medicare)	<5	5-8	9-13	14-21	>21
Esophagectomy					
No. of Medicare patients	861	1,817	1,091	1,393	1,175
No. of hospitals	618	620	187	119	31
Average no. of procedures/yr (49% Medicare)	<2	2-4	5-7	8-19	>19
Pancreatic resection					
No. of Medicare patients	1,563	2,757	1,885	2,166	2,159
No. of hospitals	1,028	555	168	90	27
Average no. of procedures/yr (50% Medicare)	<1	1-2	3-5	6-16	>16
Nephrectomy					
No. of Medicare patients	10,732	11,724	11,899	12,385	12,250
No. of hospitals	1,916	606	380	253	137
Average no. of procedures/yr (43% Medicare)	<7	7-12	13-19	20-31	>31
Cystectomy					
No. of Medicare patients	4,262	5,380	3,149	4,833	4,725
No. of hospitals	1,438	550	190	180	64
Average no. of procedures/yr (63% Medicare)	<2	2-3	4-5	6-11	>11
Pulmonary resection					
Lobectomy					
No. of Medicare patients	14,816	15,731	14,759	15,469	14,788
No. of hospitals	1,806	456	248	164	79
Pneumonectomy					
No. of Medicare patients	1,969	2,098	2,072	2,088	2,183
No. of hospitals	944	445	246	163	79
Average no. of procedures/yr (56% Medicare)	<9	9-17	18-27	28-46	>46

*To estimate the average total hospital volume, we divided the observed Medicare volume (per year) by the proportion of Medicare patients undergoing each procedure (as determined on the basis of the Nationwide Inpatient Sample).

TABLE 2. CHARACTERISTICS OF THE PATIENTS ACCORDING TO HOSPITAL VOLUME.*

PROCEDURE	HOSPITAL VOLUME				
	VERY LOW	LOW	MEDIUM	HIGH	VERY HIGH
	percent				
Coronary-artery bypass grafting					
Age >75 yr	35.8	36.2	36.1	36.6	36.8
Female sex	35.2	34.7	35.1	35.3	35.0
Black race	4.7	3.5	3.5	3.6	3.4
Nonelective admission	60.8	60.4	59.8	60.9	60.8
Charlson score ≥3	9.2	9.0	9.1	9.5	10.1
Aortic-valve replacement					
Age >75 yr	49.6	50.3	51.0	53.1	52.3
Female sex	43.9	42.6	43.1	44.9	42.5
Black race	4.4	4.0	3.2	3.5	2.0
Nonelective admission	44.5	43.8	43.8	43.9	40.8
Charlson score ≥3	8.5	8.7	8.7	10.0	9.4
Mitral-valve replacement					
Age >75 yr	39.6	41.1	40.9	43.3	42.4
Female sex	59.0	58.2	57.7	59.4	58.1
Black race	4.7	4.2	3.4	3.8	2.6
Nonelective admission	49.4	49.4	49.7	48.1	45.6
Charlson score ≥3	9.4	9.5	10.1	10.7	10.9
Carotid endarterectomy					
Age >75 yr	49.0	48.8	47.3	46.4	46.3
Female sex	45.8	44.2	43.4	43.5	42.2
Black race	4.0	2.6	2.6	2.6	2.4
Nonelective admission	36.0	32.1	33.5	31.1	30.3
Charlson score ≥3	10.0	10.1	9.8	9.8	10.2
Lower-extremity bypass					
Age >75 yr	49.5	50.6	50.4	50.4	50.8
Female sex	46.6	46.8	46.1	45.5	46.1
Black race	12.4	10.5	12.6	10.5	12.6
Nonelective admission	44.6	42.2	41.1	37.2	38.4
Charlson score ≥3	20.1	20.0	19.6	20.8	21.0
Elective repair of abdominal aortic aneurysm					
Age >75 yr	44.4	43.8	42.5	42.6	42.2
Female sex	25.1	23.3	23.0	22.2	22.3
Black race	4.2	2.5	2.4	2.5	2.4
Nonelective admission	31.8	26.3	27.3	25.3	25.5
Charlson score ≥3	9.1	9.2	9.5	9.7	10.1
Colectomy					
Age >75 yr	60.9	60.5	59.9	59.7	59.3
Female sex	56.2	56.1	55.9	56.0	54.6
Black race	8.8	7.3	7.6	6.9	7.2
Nonelective admission	55.7	47.4	48.5	44.4	43.0
Charlson score ≥3	41.1	44.3	45.3	45.1	45.9
Gastrectomy					
Age >75 yr	56.8	54.9	54.0	53.3	48.9
Female sex	44.0	41.3	41.9	39.7	37.4
Black race	13.7	12.9	13.2	11.3	11.3
Nonelective admission	51.4	44.7	41.9	41.0	29.7
Charlson score ≥3	56.2	58.8	59.3	60.3	62.3
Esophagectomy					
Age >75 yr	31.9	31.7	31.3	28.6	28.1
Female sex	22.8	24.0	24.6	23.4	25.8
Black race	12.0	7.6	6.7	6.5	5.5
Nonelective admission	37.4	26.9	20.5	20.5	7.6
Charlson score ≥3	43.3	43.3	45.4	41.6	43.0
Pancreatic resection					
Age >75 yr	37.7	36.3	35.3	36.8	35.2
Female sex	48.0	51.3	49.0	48.9	47.4
Black race	8.1	8.0	6.8	6.0	5.0
Nonelective admission	50.6	43.7	43.9	35.4	18.6
Charlson score ≥3	46.4	51.7	53.4	55.6	64.5
Nephrectomy					
Age >75 yr	46.1	45.8	44.3	43.2	41.6
Female sex	44.9	44.0	42.8	42.9	41.4
Black race	7.7	5.2	5.5	6.0	6.5
Nonelective admission	35.4	28.5	24.7	23.4	20.2
Charlson score ≥3	19.2	19.9	20.2	21.1	23.1

TABLE 2. CONTINUED.

PROCEDURE	HOSPITAL VOLUME				
	VERY LOW	LOW	MEDIUM	HIGH	VERY HIGH
	percent				
Cystectomy					
Age >75 yr	41.1	42.7	44.0	43.3	39.2
Female sex	23.3	22.5	20.8	20.8	18.3
Black race	6.6	3.5	3.1	3.3	3.8
Nonelective admission	27.4	22.9	18.4	17.0	14.9
Charlson score ≥ 3	34.0	33.8	35.6	36.6	43.1
Pulmonary lobectomy					
Age >75 yr	33.8	34.7	33.1	33.7	32.0
Female sex	43.1	44.3	43.0	43.6	44.2
Black race	7.3	4.2	4.3	4.5	4.7
Nonelective admission	26.1	21.6	19.3	14.3	12.7
Charlson score ≥ 3	31.1	32.0	32.1	34.1	37.0
Pneumonectomy					
Age >75 yr	25.9	26.0	24.5	24.5	20.8
Female sex	28.2	27.0	28.1	27.5	27.3
Black race	8.1	3.9	4.0	4.8	5.0
Nonelective admission	28.2	24.0	24.6	17.6	15.8
Charlson score ≥ 3	52.4	54.3	56.8	57.5	59.9

*The Charlson score is a measure of the number of coexisting conditions, weighted according to their relative effects on mortality.

in the quality of surgery between high-volume and low-volume hospitals. First, the effect is large. For some procedures, mortality at low-volume centers was several times as high as at high-volume hospitals — a difference that is too great to be attributed to chance or unmeasured confounding. Second, relations between volume and outcome are remarkably consistent over time and across studies. According to one recent structured review of the literature, 123 of 128 analyses involving 40 different procedures (96 percent) found lower mortality at high-volume hospitals (differences were statistically significant in 80 percent of these analyses).⁵ Only 4 of the 128 (3 percent) found higher mortality rates at high-volume hospitals, but none of these findings were statistically significant. And finally, the link between surgical volume and mortality is clinically plausible. Although the mechanisms underlying the relations between volume and outcome have not been fully characterized, high-volume hospitals may have more surgeons who specialize in specific procedures, more consistent processes for postoperative care, better-staffed intensive care units, and greater resources, in general, for dealing with postoperative complications.

Our analysis has several limitations. First, because we studied only Medicare patients, our results may not be generalizable to patients under 65 years of age. However, there is no evidence that age affects the relations between volume and outcome. Second, our measure of volume was imperfect. We estimated total

hospital volume by extrapolating from Medicare volume, not by direct measurement. Although Medicare and total volumes are highly correlated at the hospital level, there probably remains some degree of misclassification of hospital-volume status, which would tend to bias our analysis toward the null hypothesis (no effect of volume on outcome). Third, because our primary goal was to estimate the potential effect of referral policies that focus exclusively on volume, we did not attempt to adjust for characteristics of the provider that are likely to be highly correlated with volume. Analyses that aimed to assess the independent effect of hospital volume would need to account for other variables that may influence mortality, including hospital size and teaching status, the volume of procedures performed by a particular surgeon, and staffing patterns in the intensive care unit.²⁴⁻²⁷

Finally, because we relied on administrative data, we may not have accounted adequately for differences in case mix among strata of hospital volume. Administrative data are limited in their ability to differentiate patients according to the severity of illness.^{21,22,28,29} Age and the prevalence of coexisting conditions did not vary substantially according to hospital volume in our data set. However, even for conditions for which the procedure itself is almost always elective, patients at lower-volume hospitals were more likely to have been admitted nonelectively. Conversely, patients at higher-volume hospitals were more likely to have had recent nonelective admissions elsewhere. Although these find-

TABLE 3. OPERATIVE MORTALITY RATES AND THEIR ASSOCIATION WITH HOSPITAL VOLUME.*

PROCEDURE	HOSPITAL VOLUME				
	VERY LOW	LOW	MEDIUM	HIGH	VERY HIGH
Coronary-artery bypass grafting					
Observed mortality rate (%)	6.1	5.5	5.3	5.1	4.8
Unadjusted odds ratio (95% CI)	1.0	0.91 (0.86–0.97)	0.88 (0.82–0.93)	0.84 (0.79–0.90)	0.79 (0.73–0.86)
Adjusted odds ratio (95% CI)	1.0	0.92 (0.86–0.98)	0.89 (0.83–0.95)	0.84 (0.78–0.90)	0.79 (0.73–0.86)
Aortic-valve replacement					
Observed mortality rate (%)	9.9	9.2	9.1	8.7	7.6
Unadjusted odds ratio (95% CI)	1.0	0.91 (0.84–0.99)	0.91 (0.84–0.99)	0.87 (0.80–0.95)	0.75 (0.65–0.85)
Adjusted odds ratio (95% CI)	1.0	0.92 (0.85–0.99)	0.91 (0.84–0.99)	0.84 (0.77–0.92)	0.75 (0.66–0.86)
Mitral-valve replacement					
Observed mortality rate (%)	16.1	15.0	14.4	13.4	12.5
Unadjusted odds ratio (95% CI)	1.0	0.92 (0.84–1.01)	0.87 (0.79–0.97)	0.81 (0.73–0.90)	0.74 (0.66–0.84)
Adjusted odds ratio (95% CI)	1.0	0.91 (0.83–1.00)	0.86 (0.78–0.96)	0.79 (0.71–0.88)	0.74 (0.65–0.84)
Carotid endarterectomy					
Observed mortality rate (%)	2.0	1.9	1.8	1.7	1.7
Unadjusted odds ratio (95% CI)	1.0	0.91 (0.85–0.98)	0.88 (0.81–0.95)	0.82 (0.76–0.89)	0.82 (0.75–0.90)
Adjusted odds ratio (95% CI)	1.0	0.95 (0.88–1.02)	0.91 (0.84–0.99)	0.88 (0.81–0.95)	0.88 (0.80–0.96)
Lower-extremity bypass					
Observed mortality rate (%)	6.1	5.8	5.5	5.5	4.9
Unadjusted odds ratio (95% CI)	1.0	0.94 (0.88–0.99)	0.89 (0.83–0.94)	0.90 (0.84–0.96)	0.78 (0.72–0.85)
Adjusted odds ratio (95% CI)	1.0	0.94 (0.89–1.00)	0.90 (0.85–0.97)	0.94 (0.87–1.01)	0.81 (0.74–0.88)
Elective repair of abdominal aortic aneurysm					
Observed mortality rate (%)	7.8	5.9	5.2	5.3	4.4
Unadjusted odds ratio (95% CI)	1.0	0.75 (0.69–0.81)	0.66 (0.60–0.72)	0.66 (0.61–0.73)	0.54 (0.49–0.60)
Adjusted odds ratio (95% CI)	1.0	0.79 (0.73–0.86)	0.70 (0.64–0.76)	0.71 (0.65–0.78)	0.58 (0.53–0.65)
Colectomy					
Observed mortality rate (%)	7.4	6.9	6.4	6.1	5.4
Unadjusted odds ratio (95% CI)	1.0	0.93 (0.88–0.98)	0.86 (0.81–0.90)	0.81 (0.77–0.86)	0.73 (0.68–0.77)
Adjusted odds ratio (95% CI)	1.0	0.98 (0.93–1.03)	0.89 (0.84–0.94)	0.89 (0.84–0.93)	0.80 (0.76–0.85)
Gastrectomy					
Observed mortality rate (%)	13.0	12.7	11.1	11.3	8.7
Unadjusted odds ratio (95% CI)	1.0	0.98 (0.88–1.10)	0.84 (0.75–0.93)	0.85 (0.76–0.96)	0.64 (0.55–0.74)
Adjusted odds ratio (95% CI)	1.0	1.01 (0.90–1.13)	0.88 (0.79–0.99)	0.90 (0.80–1.01)	0.72 (0.63–0.83)
Esophagectomy					
Observed mortality rate (%)	23.1	18.9	16.9	11.7	8.1
Unadjusted odds ratio (95% CI)	1.0	0.78 (0.63–0.95)	0.68 (0.54–0.86)	0.44 (0.35–0.55)	0.29 (0.21–0.40)
Adjusted odds ratio (95% CI)	1.0	0.85 (0.69–1.05)	0.76 (0.60–0.97)	0.51 (0.40–0.64)	0.36 (0.26–0.50)
Pancreatic resection					
Observed mortality rate (%)	17.6	15.4	11.6	7.5	3.8
Unadjusted odds ratio (95% CI)	1.0	0.85 (0.72–1.01)	0.62 (0.50–0.76)	0.38 (0.31–0.47)	0.18 (0.13–0.26)
Adjusted odds ratio (95% CI)	1.0	0.88 (0.74–1.05)	0.64 (0.51–0.79)	0.40 (0.32–0.50)	0.20 (0.14–0.29)
Nephrectomy					
Observed mortality rate (%)	3.6	3.2	2.7	2.9	2.6
Unadjusted odds ratio (95% CI)	1.0	0.88 (0.76–1.02)	0.75 (0.64–0.87)	0.80 (0.69–0.93)	0.70 (0.59–0.82)
Adjusted odds ratio (95% CI)	1.0	0.94 (0.81–1.09)	0.83 (0.71–0.97)	0.90 (0.77–1.05)	0.80 (0.68–0.95)
Cystectomy					
Observed mortality rate (%)	6.4	6.3	4.9	4.5	2.9
Unadjusted odds ratio (95% CI)	1.0	0.99 (0.84–1.17)	0.76 (0.62–0.93)	0.69 (0.57–0.84)	0.44 (0.35–0.54)
Adjusted odds ratio (95% CI)	1.0	1.00 (0.84–1.18)	0.78 (0.64–0.96)	0.71 (0.58–0.86)	0.46 (0.37–0.58)
Pulmonary lobectomy					
Observed mortality rate (%)	6.4	5.9	5.5	5.3	4.2
Unadjusted odds ratio (95% CI)	1.0	0.91 (0.83–1.01)	0.85 (0.76–0.94)	0.81 (0.73–0.90)	0.64 (0.55–0.75)
Adjusted odds ratio (95% CI)	1.0	0.94 (0.85–1.04)	0.89 (0.80–0.99)	0.87 (0.78–0.97)	0.70 (0.60–0.81)
Pneumonectomy					
Observed mortality rate (%)	17.0	15.4	15.7	15.0	10.6
Unadjusted odds ratio (95% CI)	1.0	0.89 (0.75–1.06)	0.91 (0.77–1.08)	0.86 (0.73–1.02)	0.58 (0.47–0.73)
Adjusted odds ratio (95% CI)	1.0	0.91 (0.76–1.08)	0.93 (0.78–1.11)	0.91 (0.76–1.08)	0.62 (0.50–0.77)

*Odds ratios for death (according to quintile of volume) were adjusted for age, sex, race, year of procedure, Social Security income, urgency of admission, and Charlson score, a measure of coexisting conditions. CI denotes confidence interval.

ings raise the possibility of unmeasured differences in case mix among hospitals, we do not believe that confounding is a likely explanation for our main findings.

Although relations between volume and outcome have long been recognized, large-scale efforts to re-

duce surgical mortality by concentrating selected procedures in high-volume hospitals are only now beginning to gain momentum. The most visible of these efforts is being directed by the nonprofit Leapfrog Group, a coalition of more than 80 large public and

HOSPITAL VOLUME AND SURGICAL MORTALITY

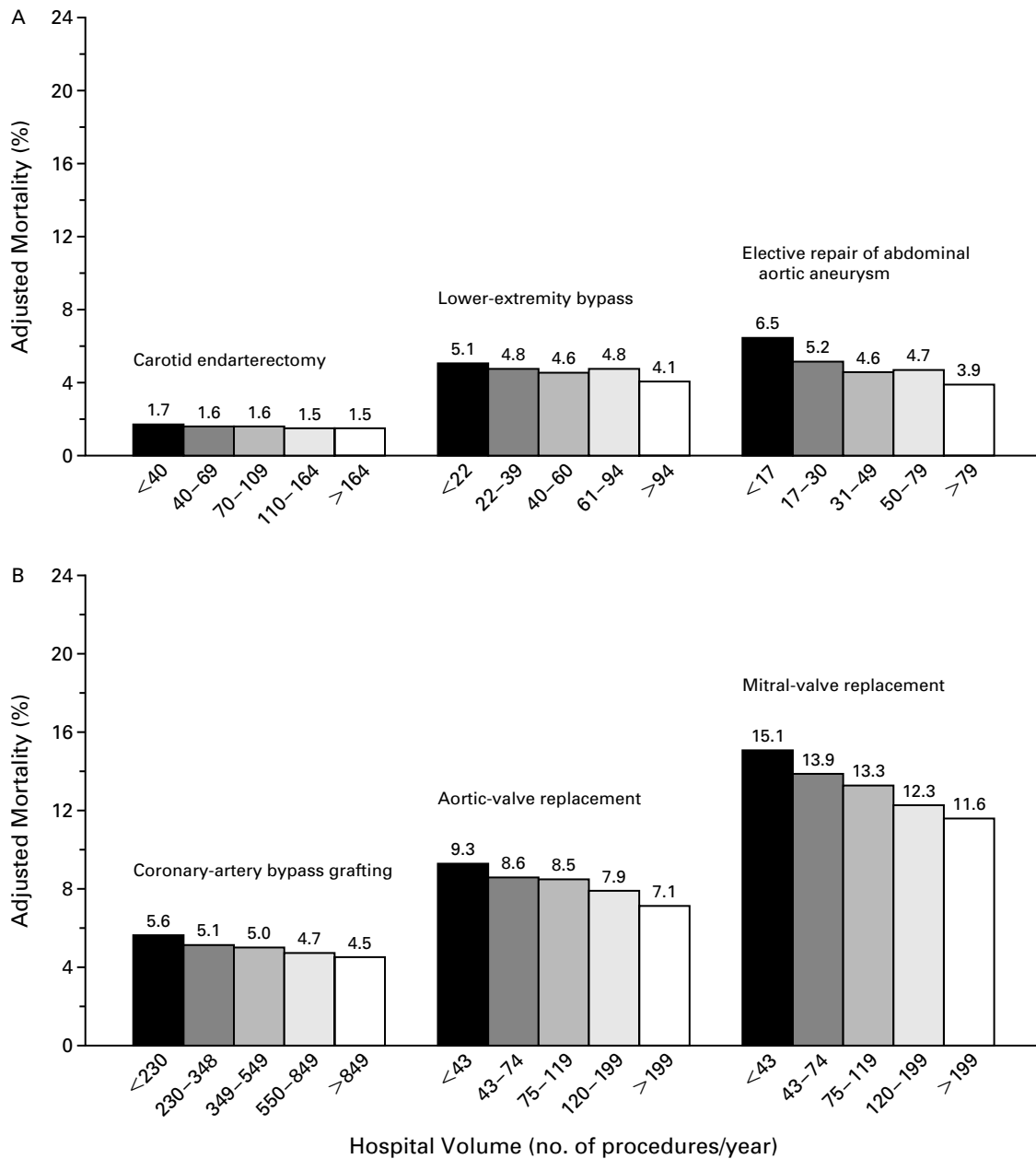


Figure 1. Adjusted In-Hospital or 30-Day Mortality among Medicare Patients (1994 through 1999), According to Quintile of Total Hospital Volume for Peripheral Vascular Procedures (Panel A) and Cardiac Procedures (Panel B). P<0.001 for all procedures. The outcomes for aortic-valve and mitral-valve replacement were stratified according to the total volume of heart-valve replacements. Values above the bars are the percent mortality.

private purchasers that insure more than 25 million persons. The coalition is encouraging both patients and payers to select hospitals that meet minimal volume standards for coronary-artery bypass surgery (500 procedures per year), coronary angioplasty (400 per year), carotid endarterectomy (100 per year), repair of abdominal aortic aneurysm (30 per year), and esoph-

agectomy for cancer (6 per year). Although our analysis does not indicate that these specific volume thresholds are better than other alternatives, it does confirm that the proposed standards could reduce the surgical mortality associated with several of these procedures.

Many may object to such initiatives aimed at concentrating selected surgical procedures in high-volume

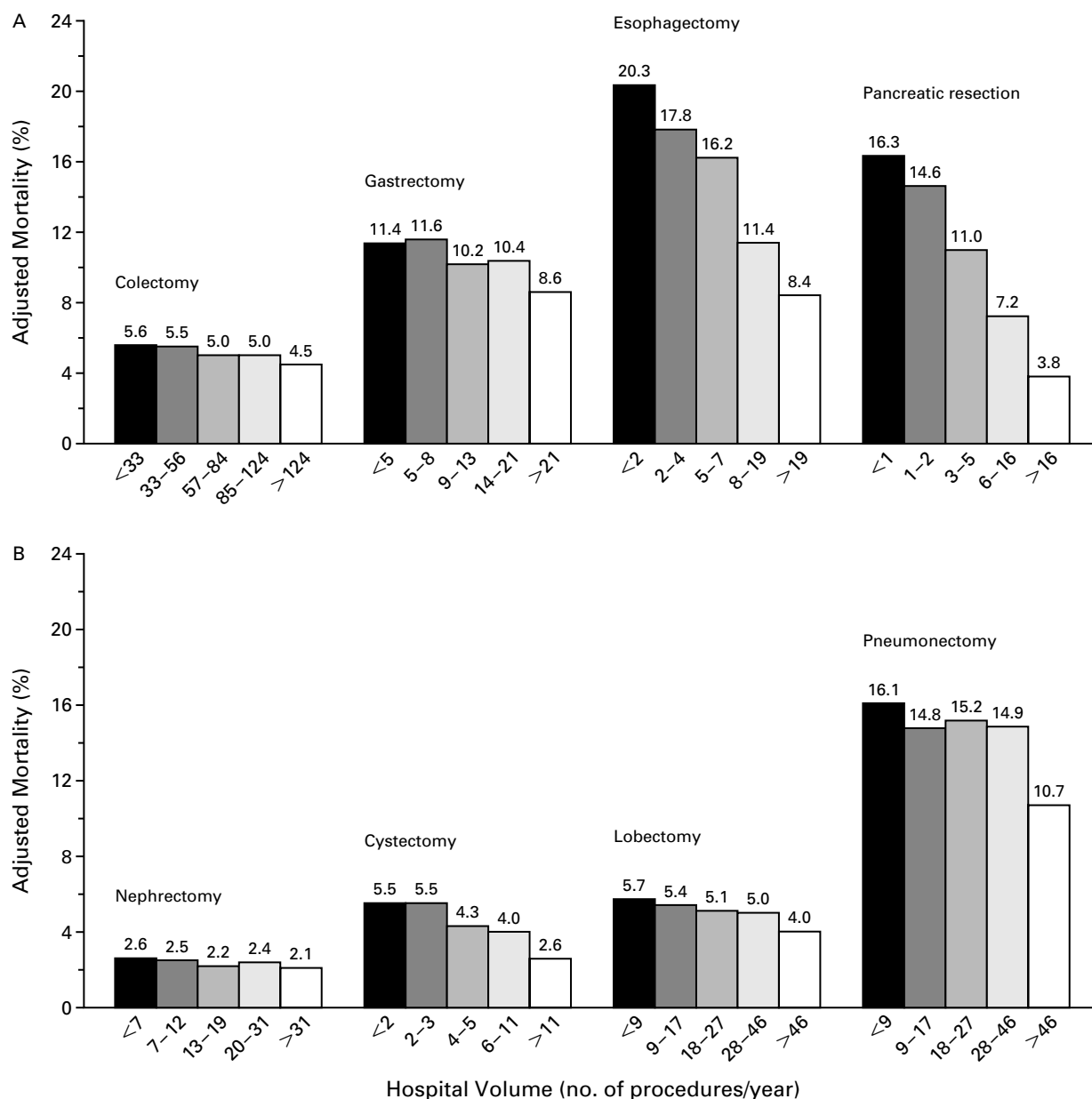


Figure 2. Adjusted In-Hospital or 30-Day Mortality among Medicare Patients (1994 through 1999), According to Quintile of Total Hospital Volume for Resections of Gastrointestinal Cancer (Panel A) and Resections of Other Cancers (Panel B). $P < 0.001$ for all procedures. The outcomes for lobectomy and pneumonectomy were stratified according to the total volume of lung resections. Values above the bars are the percent mortality.

hospitals. They may rightly point out that procedure volume is an imperfect proxy for quality — that some low-volume hospitals have excellent outcomes, whereas some high-volume hospitals have poor outcomes. Unfortunately, most patients facing high-risk surgery have no way of knowing the relative quality of the hospitals near them. Although several states currently have

public reporting systems in place,^{30,31} these efforts are largely restricted to reporting on cardiac surgery. Most other procedures are not performed frequently enough to allow assessment of procedure-specific mortality at the level of the individual hospital. Thus, in the absence of better information about surgical quality, patients undergoing many types of procedures can sub-

stantially improve their odds of survival by selecting a high-volume hospital near them.

Supported by a grant (R01 HS10141-01) from the Agency for Healthcare Research and Quality. Dr. Birkmeyer is also supported by a Career Development Award from the Veterans Affairs Health Services Research and Development program. The views expressed herein do not necessarily represent the views of the Department of Veterans Affairs or the United States Government.

REFERENCES

- 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.
- Hannan EL, O'Donnell JF, Kilburn H Jr, Bernard HR, Yazici A. Investigation of the relationship between volume and mortality for surgical procedures performed in New York State hospitals. *JAMA* 1989;262:503-10.
- Begg CB, Cramer LD, Hoskins WJ, Brennan MF. Impact of hospital volume on operative mortality for major cancer surgery. *JAMA* 1998;280:1747-51.
- Flood AB, Scott WR, Ewy W. Does practice make perfect? I. The relation between hospital volume and outcomes for selected diagnostic categories. *Med Care* 1984;22:98-114.
- Dudley RA, Johansen KL, Brand R, Rennie DJ, Milstein A. Selective referral to high-volume hospitals: estimating potentially avoidable deaths. *JAMA* 2000;283:1159-66.
- Halm EA, Lee C, Chassin MR. How is volume related to quality in health care? A systematic review of the research literature. Washington, D.C.: Institute of Medicine, 2000.
- Birkmeyer JD, Lucas FL, Wennberg DE. Potential benefits of regionalizing major surgery in Medicare patients. *Eff Clin Pract* 1999;2:277-83.
- Birkmeyer JD, Finlayson EVA, Birkmeyer CM. Volume standards for high-risk surgical procedures: potential benefits of the Leapfrog initiative. *Surgery* 2001;130:415-22.
- Marsa L. Not every hospital is created equal. *Los Angeles Times*. September 25, 2000:S1.
- Comarow A. Higher volume, fewer deaths. *U.S. News & World Report*. July 17, 2000:68-70.
- Russell TR. Volume standards for high-risk questions: an American College of Surgeons' view. *Surgery* 2001;130:423-4.
- Khuri SF. Surgeons, not General Motors, should set standards for surgical care. *Surgery* 2001;130:429-31.
- Ghali WA, Ash AS, Hall RE, Moskowitz MA. Statewide quality improvement initiatives and mortality after cardiac surgery. *JAMA* 1997;277:379-82.
- Katz DJ, Stanley JC, Zelenock GB. Operative mortality rates for intact and ruptured abdominal aortic aneurysms in Michigan: an eleven-year statewide experience. *J Vasc Surg* 1994;19:804-17.
- Finlayson EVA, Birkmeyer JD. Operative mortality with elective surgery in older adults. *Eff Clin Pract* 2001;4:172-7. [Erratum, *Eff Clin Pract* 2001;4:235.]
- Department of Health and Human Services. The international classification of diseases, 9th rev., clinical modification: ICD-9-CM. Washington, D.C.: Government Printing Office, 1998.
- Laks MP, Cohen T, Hack R. Volume of procedures at transplantation centers and mortality after liver transplantation. *N Engl J Med* 2000;342:1527.
- McCullagh P, Nelder JA. Generalized linear models. 2nd ed. New York: Chapman & Hall, 1989.
- Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40:373-83.
- Romano PS, Roos LL, Jollis JG. Adapting a clinical comorbidity index for use with ICD-9-CM administrative data: differing perspectives. *J Clin Epidemiol* 1993;46:1075-9.
- Iezzoni LI, Foley SM, Daley J, Hughes J, Fisher ES, Heeren T. Comorbidities, complications, and coding bias: does the number of diagnosis codes matter in predicting in-hospital mortality? *JAMA* 1992;267:2197-203.
- Jencks SE, Williams DK, Kay TL. Assessing hospital-associated deaths from discharge data: the role of length of stay and comorbidities. *JAMA* 1988;260:2240-6.
- Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 1986;42:121-30.
- Pronovost PJ, Jenckes MW, Dorman T, et al. Organizational characteristics of intensive care units related to outcomes of abdominal aortic surgery. *JAMA* 1999;281:1310-7.
- Ghorra S, Reinert SE, Cioffi W, Buczeko G, Simms HH. Analysis of the effect of conversion from open to closed surgical intensive care unit. *Ann Surg* 1999;229:163-71.
- Taylor DH Jr, Whellan DJ, Sloan FA. Effects of admission to a teaching hospital on the cost and quality of care for Medicare beneficiaries. *N Engl J Med* 1999;340:293-9.
- Kelly JV, Hellinger FJ. Physician and hospital factors associated with mortality of surgical patients. *Med Care* 1986;24:785-800.
- Fisher ES, Whaley FS, Krushat WM, et al. The accuracy of Medicare's hospital claims data: progress has been made, but problems remain. *Am J Public Health* 1992;82:243-8.
- Hsia DC, Krushat WM, Fagan AB, Tebbutt JA, Kusserow RP. Accuracy of diagnostic coding for Medicare patients under the prospective-payment system. *N Engl J Med* 1988;318:352-5. [Erratum, *N Engl J Med* 1990;322:1540.]
- Schneider EC, Epstein AM. Use of public performance data reports: a survey of patients undergoing cardiac surgery. *JAMA* 1998;279:1638-42.
- Hannan EL, Kumar D, Racz M, Siu AL, Chassin MR. New York State's Cardiac Surgery Reporting System: four years later. *Ann Thorac Surg* 1994;58:1852-7.

Copyright © 2002 Massachusetts Medical Society.