

Outcomes of surgery among the Medicare aged: Surgical volume and mortality

by Gerald Riley and James Lubitz

We examined the relation between surgical volume and mortality, within 60 days of surgery, for eight procedures on aged Medicare beneficiaries. Logistic regression revealed that high surgical volume was significantly associated with lower mortality for resection of the intestine, coronary artery bypass, transurethral resection of the prostate (TURP), and hip arthroplasty (excluding total hip replacement). For

cholecystectomy, total hip replacement, inguinal hernia repair, and femur fracture reduction, no relationship was found between surgical volume and postsurgical mortality. The analyses were repeated using inhospital deaths as the dependent variable, and the results indicated a considerably stronger association between volume and mortality.

Introduction

Surgical procedures performed on the aged account for a high percentage of all surgeries in the United States. A National Center for Health Statistics (NCHS) report (Pokras, 1983) estimates that in 1979 more than 20 percent of all procedures were for patients 65 years of age or over. The rate of all listed procedures per 10,000 population was estimated by NCHS to be 2,583 for the aged and only 1,371 for all ages in that year. Lubitz and Deacon (1982) have noted that the incidence of surgery among the Nation's aged increased at a faster rate between 1967 and 1977 than for younger people. Because mortality following surgery generally increases with age (Lubitz, Riley, and Newton, 1985; Jensen and Tondevold, 1979; and Gersh et al., 1983), postsurgical mortality among the aged is an important aspect of surgical outcomes.

Among hospitals, surgical outcomes can vary considerably in a manner not fully explained by measurable differences in individual patient characteristics. The National Halothane Study (National Academy of Sciences, 1966) examined the effects of halothane anesthesia on patients undergoing surgery. The study, adjusting for differences in procedures, age, and preoperative physical status, produced evidence of considerable variation in inhospital death rates among the 34 hospitals. The Stanford Center for Health Care Research (1974), using Professional Activities Study (PAS) data collected from 1,224 hospitals, studied differences in institutional mortality rates after indirectly standardizing for demographic and clinical differences in patient populations. The authors found significant differences in mortality rates among institutions. The study also addressed the relationship between various hospital characteristics (e.g., size, teaching programs, and staffing levels) and standardized mortality ratios (MR's) for various categories of procedures. The investigators found that standardized MR's differed according to hospital

characteristics, but the nature and strength of the relationship varied strongly by procedure. In general, the amount of variation in standardized MR's explained by hospital characteristics was rather small for each procedure.

Schumacher and Horn (1978) examined postoperative death rates among Maryland hospitals both for those located in standard metropolitan statistical areas (SMSA's) and those in non-SMSA's. Controlling for case mix, the investigators found an inverse relationship between cost per case and mortality rates in SMSA's, but did not find significant relationships between mortality rates and other hospital characteristics, such as number of residency training programs, hospital salary level, and presence of high technology facilities (e.g., coronary care units and intensive care units). The Commission on Professional and Hospital Activities (1970), using Professional Activity Study (PAS) data, found that after adjusting for differences in patient mix, cholecystectomy patients discharged from small hospitals experienced a much higher mortality rate than those discharged from large hospitals. Hospital size was defined in terms of total number of discharges.

Luft, Bunker, and Enthoven (1979) and Luft (1980) examined the issue of whether hospitals that perform large numbers of specific surgical procedures have different mortality outcomes than hospitals that perform relatively few of these procedures. From PAS data, hospital-specific mortality rates were compared with expected mortality rates, adjusting for age, sex, and presence of single or multiple diagnoses. Mortality rates decreased as volume of the specific operations increased for several procedures, including open heart surgery, coronary artery bypass graft, transurethral resection of the prostate (TURP), and hip arthroplasty (total replacement). This association was determined to exist independent of the effects of hospital size such as total admissions and total number of operations performed. No significant volume-mortality relationship was found for some procedures, e.g., vagotomy. Although the data suggest that greater surgical experience may lead to lower mortality rates for some procedures, Luft (1980) also provided evidence that suggests a referral effect may exist, i.e., that better outcomes may lead

Reprint requests: Gerald Riley, Health Care Financing Administration, 2-C-15 Oak Meadows Building, 6325 Security Boulevard, Baltimore, MD 21207.

to more referrals and higher volume. Regardless of the exact nature of the relationship between volume and mortality, Luft, Bunker, and Enthoven (1979) have argued that the data support the concept of establishing regional referral centers for some procedures.

Additional evidence supporting the establishment of regional centers for performing certain procedures was reported recently by Flood, Scott, and Ewy (1984a and 1984b). They used 1972 data from the Commission on Professional and Hospital Activities, and found strong and consistent evidence that high volume is associated with lower in-hospital mortality for surgical patients, after controlling for patient mix. In particular, low-volume hospitals were associated with poorest outcomes relative to expectations for low-risk surgical patients.

Our study examines the relationship between hospitals' postsurgical mortality rates and surgical volume for aged Medicare beneficiaries undergoing eight types of procedures in 1979 or 1980. The procedures are TURP, hip arthroplasty (total hip replacement and other arthroplasties are examined separately), femur fracture reduction,¹ resection of the intestine, cholecystectomy, inguinal hernia repair, and coronary artery bypass.

These procedures were chosen for several reasons. All, with the exception of coronary artery bypass surgery, are common among the aged. For five of the eight study procedures, more than half of all operations were performed on the aged. In 1979, 86 percent of all hip arthroplasties (not including total hip replacement), 74 percent of all TURP's, 69 percent of all femur fracture reductions, 57 percent of total hip replacements, and 52 percent of resections of the intestine were performed on the aged (National Center for Health Statistics, 1982). Between 23 percent and 28 percent of the coronary artery bypass surgeries, repairs of inguinal hernia, and cholecystectomies were performed on aged patients.

In addition, all these procedures, except for femur fracture reduction, are often of a non-emergency nature. This means that, aside from the decision to operate or not, there is often considerable discretion available as to which hospital the procedure will be performed in. Lastly, several of the procedures exhibit rather high mortality rates following surgery (Lubitz, Riley, and Newton, 1985). The association of volume and mortality rates is particularly important for the aged, because the relative frailty of this population and the frequent presence of co-morbidity may make the aged more sensitive to the effects of surgical experience.

¹ Although the *International Classification of Diseases, Ninth Revision, Clinical Modifications (ICD-9-CM)* surgical codes do not identify the specific part of the femur that is fractured, for elderly patients the majority of breaks occur at the neck of the femur. In this study, among patients for whom the principal diagnosis was a fractured femur, more than 90 percent of the cases had a diagnosis of fracture at the neck of the femur.

Data and methods

Data for this study were obtained from two files of the Medicare Statistical System of the Health Care Financing Administration (HCFA): the Medicare provider analysis and review (MEDPAR) file and the enrollment file. The MEDPAR file contains information on discharges from short-stay hospitals for a 20-percent probability sample of Medicare beneficiaries (selected on the basis of their Medicare identification numbers). The information includes demographic data (e.g., age, sex), hospital data (e.g., size, location), and stay data (e.g., principal surgical procedure, principal diagnosis, and in-hospital patient death). In 1979 and 1980, the principal surgical procedure and principal diagnosis were coded for each discharge using the *International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM)*. The records of hospital stays for the eight selected procedures were linked to the enrollment file to identify all deaths, regardless of where the death occurred, for 60 days following surgery.

One hospital-specific variable was computed from the MEDPAR file. This variable, referred to in this article as VOLUME, is the annual number of surgeries of a given type performed on the Medicare aged in a specific hospital. VOLUME was derived by summing, for each hospital, the number of surgeries performed in the appropriate year and then multiplying by 5.

Several limitations are present in the data. First, the experience of many hospitals with a very low volume of specific surgeries is not represented in the analysis. This is true because, under the 20-percent sample selection criterion, no patients are selected from many hospitals that perform only a few operations of a specific type in a given year. Consequently, it should be more difficult to detect mortality-volume relationships that exist at a low range of surgical volumes. A related problem is that some error is present in the measurement of surgical volumes for individual hospitals because the MEDPAR file discharges are selected as a national sample and not on a hospital-by-hospital basis. This error can be substantial at very low sample volumes (Table 1). Second, the exact date of death is not always identified because, in the Medicare enrollment file, the date of death is sometimes coded as the last day of the month of death. This is most often true for railroad retirees, who constituted approximately 3 percent of our sample. Railroad retirees were therefore excluded from the analyses, and for the remainder of the sample it was estimated that 5 percent of all deaths were erroneously recorded as having occurred on the last day of the month. The effect of erroneously recorded death dates would be to slightly underestimate the number of deaths occurring within 60 days of surgery, but there is no reason to believe that this invalidates the results relating to mortality-volume relationships.

Third, there are known to be some problems with the reliability of diagnostic and surgical data in the

Table 1
Approximate standard errors associated with estimates of the surgical volume at an individual hospital, by sample sizes¹

Sample surgeries (individual hospital)	Population estimate	Approximate standard error
3	15	7.75
6	30	10.95
10	50	14.14
15	75	17.31
25	125	22.35
40	200	28.26

¹Taken from the 20-percent national Medicare provider analysis and review (MEDPAR) file.

SOURCE: Health Care Financing Administration, Bureau of Data Management and Strategy: Data from the Medicare Statistical System, 1979-80.

Medicare statistical system (National Academy of Sciences, 1977). The problems uncovered in the National Academy of Sciences study are described in another article (Lubitz, Riley, and Newton, 1985). The study did not reveal any errors that would appear to systematically bias our study results.²

For six of the procedures, cases with a principal diagnosis of cancer for the stay in which the study procedure was performed were excluded from the sample; the six procedures are hip arthroplasty (total), hip arthroplasty (other), femur fracture reduction, inguinal hernia repair, cholecystectomy, and coronary artery bypass. Among beneficiaries undergoing these operations, there were few cancer patients (1.2 percent); their mortality experience was considerably different from that of the noncancer patients. Cancer patients were retained in the sample for TURP and resection of the intestine because there were significant numbers of such patients. The analyses control for the effect of a cancer diagnosis on mortality rates. The study also does not include Medicare beneficiaries entitled because of disability (all of whom are under 65), who constitute approximately 10 percent of the Medicare population.

The hospital discharge is the unit of analysis for this study. Discharges of aged beneficiaries with the appropriate ICD-9-CM surgical codes, occurring in 1979 and 1980, were selected and pooled across both years. If an individual underwent more than one hospitalization for the study procedures during 1979-80, each discharge was included as a separate observation. In addition to individual patient attributes for each discharge, such as age and sex, characteristics of the hospital in which the surgery was performed were appended to the unit record, including the annual number of procedure-specific surgeries from the MEDPAR file (VOLUME). Annual numbers of surgeries were computed for the year in which the discharge occurred. It should be noted that for patients undergoing either hip arthroplasty (total) or hip arthroplasty (other),

² For this study, certain consistency edits were performed to identify other kinds of miscoding. For example, 2 percent of the TURP cases were identified as having been performed on women, and were excluded from the study.

VOLUME represents the sum of the operations of either type performed in a given hospital and year.

The surgical outcome focused on in the study was whether the patient died within 60 days of surgery. This time period was chosen because of evidence that high mortality persists for these procedures for at least 2 months after surgery (Lubitz, Riley, and Newton, 1985). The use of deaths within 60 days of surgery represents an important difference from earlier studies on the relation of surgical volume to mortality; those studies were limited to in-hospital deaths (Luft, 1980; Flood et al. 1984a and 1984b). Following patients for 60 days removes any possible confounding effect of a relation between length of hospital stay and other variables. The associations between outcome and patient and hospital attributes were examined using multiple logistic regression, with the dependent variable defined as death within 60 days of surgery (1 = yes, 0 = no). Logistic regression is commonly used to ascertain the association between one or more independent variables and a binary response variable. The LOGIST procedure of statistical analysis system (SAS) was used to estimate the regression relationships (Harrell, 1983). Throughout the analyses, a significance level of .05 is used, although *p*-values are shown to indicate the strength of the relationships.

Findings

The sample sizes of the eight study procedures are shown in Table 2, along with the number of deaths that occurred within 60 days of surgery. The most common procedure was TURP with 55,742 sample cases in 1979-80, and the least common procedures were coronary artery bypass, with 6,157, and hip arthroplasty (total), with 9,862. The remainder were in the range of 17,000 to 35,000. Mortality rates varied substantially by procedure. The highest mortality rates within 60 days of surgery were experienced by patients undergoing resection of the intestine (11.0 percent), and femur fracture reduction (10.2 percent). In the latter case, the high death rate is partly because of the advanced age of the patients (mean 80.4 years), as indicated in Table 3. The lowest death rate was for patients undergoing inguinal hernia repair (1.5 percent).

The independent variables examined in the study, with their mean values by procedure, are listed in Table 3. The number of hospitals represented is also given by procedure. Average annual surgical volume for aged Medicare patients for specific procedures (VOLUME) was highest for TURP, with an annual average of 84 such operations in hospitals performing one or more TURP's. The second highest surgical volume was attributed to coronary artery bypass, with 70 such procedures per hospital. This fact and the fact that only 990 hospitals are represented in our sample for coronary artery bypass suggest that this operation is concentrated in relatively few hospitals and that the volume of coronary artery bypass procedures tends to be large in these hospitals. The

Table 2

Sample size and deaths occurring within 60 days of surgery for Medicare enrollees 65 years of age or over, by surgical procedure: 1979-80

Surgical procedure	ICD-9-CM ¹ codes	Sample size	Number of deaths within 60 days of surgery	Percent of patients dying within 60 days of surgery
Transurethral resection of the prostate	60.2	55,742	1,638	2.9
Femur fracture reduction	79.05, 79.15, 79.25, 79.35	20,161	2,057	10.2
Resection of the intestine	45.6, 45.7, 45.8	22,560	2,492	11.0
Cholecystectomy	51.2	34,693	1,468	4.2
Inguinal hernia repair	53.0, 53.1	32,721	482	1.5
Coronary artery bypass	36.1	6,157	398	6.5
Hip arthroplasty	81.5, 81.6	27,490	1,792	6.5
Total replacement	81.5	9,862	228	2.3
Other hip arthroplasty	81.6	17,628	1,564	8.9

¹ *International Classification of Diseases, Ninth Revision, Clinical Modification.*

SOURCE: Health Care Financing Administration, Bureau of Data Management and Strategy: Data from the Medicare Statistical System, 1979-80.

smallest procedure-specific surgical volumes were attributed to femur fracture reduction, with 34 mean annual operations, and resection of the intestine, with 36 mean annual operations. There is an unknown degree of upward bias in these average figures because some hospitals with a very low volume of specific procedures are not identified for the study, given the 20-percent sample size limitation of the MEDPAR file.

Coronary artery bypass differs from other procedures with respect to the hospital-specific variables. It tends to be performed in larger hospitals, as measured by bed size. More than 70 percent of coronary artery bypass operations were performed in hospitals with a medical school affiliation, compared with 52 percent for the next highest procedure, hip arthroplasty (total). The percentage of coronary artery bypass procedures performed in proprietary hospitals was somewhat lower (5 percent) than for any other procedure, and nearly all (98 percent) were performed in a hospital located in an SMSA area.

The percentage of patients dying within 60 days of surgery varied by surgical volume for several study procedures (Table 4). Mortality rates exhibited a rather consistent decrease as volume increased for resection of the intestine, coronary artery bypass, TURP, and hip arthroplasty (other). Mortality rates declined most sharply for resection of the intestine, from 12.4 percent for the lowest volume group to 9.7 percent for the highest. Hip arthroplasty (total) patients exhibited a substantial decrease in mortality among the highest volume hospitals. Repair of inguinal hernia patients exhibited a less consistent decline in mortality as volume increased, and femur fracture reduction and cholecystectomy patients did not show any consistent relation between death rate and surgical volume.

In Table 4, the percentage of procedures performed in low-volume and high-volume hospitals is shown. On the one hand, more than one-half of all operations on the Medicare aged for femur fracture reduction and resection of the intestine are done in low-volume hospitals (sample size of fewer than seven procedures). On the other hand, procedures

for TURP, coronary artery bypass, and hip arthroplasty (total) are frequently performed in high-volume hospitals (sample size of more than 15 procedures). The relative concentration of these procedures in high-volume hospitals probably reflects the electiveness of these procedures. Even so, approximately 27 percent of coronary artery bypass and hip arthroplasty (total) operations were done in low-volume hospitals.

The number of surgeries performed does not represent any hospital's complete surgical experience for the study procedures, but rather that with aged Medicare beneficiaries. Consequently, the study focuses on the relationship between a hospital's institutional experience in the surgical treatment of aged patients and the mortality of those patients. Surgical care of aged patients often requires special attention to co-morbid conditions, and involves substantially greater risk of complications and death. Thus, surgical volume related specifically to aged patients is an important measure of experience.

The relationship between mortality rates and surgical volume is more clearly identified through the multiple logistic regression equations developed separately for each procedure (Table 5). Because the regression coefficients are not readily interpretable in a logistic regression equation, only the signs of the coefficients are given in Table 5 along with their approximate *p*-values. A negative sign indicates that mortality decreases with an increase in the value of the independent variable and vice versa. The natural logarithm of VOLUME was used to represent surgical volume. (The logarithm of VOLUME was used because previous research has shown that the curve describing the association between surgical volume and mortality "flattens" as surgical volume increases [Luft, 1980]): Dummy variables were used to indicate sex, geographic region, presence of specific principal diagnoses, medical school affiliation of the hospital in which the surgery was performed, proprietary status, and location in an SMSA area. For resection of the intestine, cholecystectomy, and inguinal hernia repair, an additional dummy variable was included to indicate whether or not the date of surgery and the date of admission were the same.

Table 3

Mean values of independent variables, by procedure for Medicare enrollees 65 years of age or over: 1979-80

Independent variable	Procedure							
	Trans-urethral resection of prostate	Femur fracture reduction	Resection of the intestine	Cholecystectomy	Inguinal hernia repair	Coronary artery bypass	Hip arthroplasty (total)	Hip arthroplasty (other)
Patient-specific	Mean value							
Age	74.5	80.4	75.1	73.2	73.4	69.2	73.8	79.3
Sex ¹	1.00	.21	.42	.35	.88	.70	.34	.21
Emergency ²	—	—	.06	.03	.06	—	—	—
Diagnosis ³	¹⁰ .17	—	¹⁰ .52	—	¹² .06	—	¹³ .09	¹³ .74
	—	—	¹¹ .13	—	—	—	—	—
Geographic region:⁴								
Northeast	.22	.22	.27	.22	.23	.16	.22	.19
South	.32	.33	.29	.34	.32	.30	.22	.37
North-Central	.28	.29	.27	.28	.27	.28	.33	.28
Western	.18	.16	.17	.16	.18	.26	.23	.16
Hospital-specific								
MED ⁵	.43	.43	.46	.38	.39	.72	.52	.43
SMSA ⁶	.79	.77	.81	.73	.74	.98	.83	.79
PROP ⁷	.09	.07	.08	.09	.09	.05	.06	.08
BEDS ⁸	360	349	371	321	330	571	400	359
VOLUME ⁹	84	34	36	41	42	70	1459	1445
Number of hospitals represented	7,081	5,960	6,403	8,280	8,168	990	156061	

¹ Sex: 1 = Male; 0 = Female.

² Emergency admission is defined as one in which surgery was performed on the same day as admission to the hospital.

³ Diagnosis: 1 = Presence of indicated diagnosis; 0 = Absence.

⁴ Geographic region: 0 = Patient not residing in that region. Northeast: 1 = Patient residing in Northeast region. South: 1 = Patient residing in South region. North-Central: 1 = Patient residing in North-Central region. West: 1 = Patient residing in West region. This variable is not included in the regression.

⁵ MED is medical school affiliation: 1 = Medical school affiliation; 0 = No affiliation.

⁶ SMSA is standard metropolitan statistical area: 1 = Location in SMSA; 0 = Location other than SMSA.

⁷ PROP is proprietary: 1 = Proprietary hospital; 0 = Nonproprietary hospital.

⁸ BEDS refer to bed size.

⁹ Number of surgeries of specific type performed on the Medicare aged. Inflated from 20-percent sample of discharges.

¹⁰ 1 = Principal diagnosis of cancer; 0 = Other principal diagnosis.

¹¹ 1 = Principal diagnosis of diverticula of intestine; 0 = Other principal diagnosis.

¹² 1 = Principal diagnosis of obstructed hernia; 0 = Other principal diagnosis.

¹³ 1 = Principal diagnosis of femur fracture; 0 = Other principal diagnosis.

¹⁴ VOLUME represents number of arthroplasties performed of either type (total or other).

¹⁵ All hip arthroplasties combined.

Table 4

Crude mortality rates for Medicare enrollees 65 years of age or over, by procedure and range of volume: 1979-80

Procedure and range of volume ¹	Number of sample discharges	Percent dying within 60 days of surgery	Procedure and range of volume ¹	Number of sample discharges	Percent dying within 60 days of surgery
Transurethral resection of the prostate			Inguinal hernia repair		
Total	55,742	2.9	Total	32,721	1.5
Less than 9	15,687	3.2	Less than 5	10,934	1.5
9-14	13,951	3.0	5-7	8,221	1.8
15-22	12,213	2.8	8-11	6,355	1.4
More than 22	13,891	2.7	More than 11	7,211	1.2
Femur fracture reduction			Coronary artery bypass		
Total	20,161	10.2	Total	6,157	6.5
Less than 4	6,217	10.8	Less than 7	1,697	7.3
4-6	5,899	10.1	7-11	1,524	6.6
7-9	3,615	9.7	12-19	1,437	6.1
More than 9	4,430	10.0	More than 19	1,499	5.7
Resection of intestine			Hip arthroplasty (total)		
Total	22,560	11.0	Total	9,862	2.3
Less than 4	6,739	12.4	Less than 6	2,643	2.6
4-6	6,016	11.5	6-9	2,391	2.2
7-10	5,152	10.1	10-15	2,407	3.0
More than 10	4,653	9.7	More than 15	2,421	1.5
Cholecystectomy			Hip arthroplasty (other)		
Total	34,693	4.2	Total	17,628	8.9
Less than 5	10,513	4.1	Less than 4	4,233	10.0
5-7	8,294	4.4	4-7	5,248	8.5
8-11	7,909	4.2	8-12	4,054	9.1
More than 11	7,977	4.2	More than 12	4,093	8.1

¹ Surgical volume figures are not inflated from the 20-percent national sample.

Surgery on the day of admission may indicate that the admission was for an emergency or urgent case. As further controls for patient mix, length of stay and total charges were considered for inclusion as independent variables; however, they were dropped because they could equally well be considered dependent variables, e.g., a poor surgical outcome often involves complications that increase length of stay as well as total charges. Dummy variables indicating proprietary status and SMSA location were excluded from the equation for coronary artery bypass because so few of these procedures were performed in proprietary hospitals or non-SMSA areas.

Under the model described above, surgical volume exhibits a significant association with mortality for several of the study procedures, with higher volumes being associated with a lower probability of mortality. Resection of the intestine patients ($p < .001$) and those undergoing TURP ($p = .017$) exhibit a highly significant relationship between volume and mortality. Patients undergoing coronary artery bypass or hip arthroplasty (other) also show a significant association between volume and mortality ($p = .031$ and $p = .043$ respectively). The regression coefficients for patients undergoing femur fracture reduction, cholecystectomy, repair of inguinal hernia, and hip arthroplasty (total) are not statistically significant with respect to LOG (VOLUME). It is interesting to observe that all eight of the regression

coefficients with respect to LOG (VOLUME) are negative, which would not be expected in the absence of any volume-mortality relationships.

It cannot be determined with certainty whether the association between VOLUME and mortality reflects the effects of procedure-specific surgical volume or the volume of surgical cases in general. It is possible, for example, that general operating room experience acquired by hospital personnel may affect outcomes for certain procedures more so than experience with that particular type of procedure. If this is true, then our procedure-specific volume variable may only be a proxy for general surgical volume. We attempted to measure the effect of general surgical volume on mortality by introducing into the equations a general surgical volume variable measuring the number of all operations performed on the elderly; the high positive correlations among the general and procedure-specific volume variables and bed size (BEDS) (Table 6) produced substantial multicollinearity problems in the models, with the exception of coronary artery bypass. Thus, no firm conclusion could be drawn about the relationship between mortality and general surgical volume.

Several other variables exhibit consistently strong associations with postsurgical mortality in Table 5. Age exhibits a strong positive association with mortality for all eight procedures ($p < .001$). The regional dummy variables exhibit a consistent pattern for several procedures, with superior outcomes for

Table 5
Multiple logistic regression for specific procedures with death within 60 days as
the dependent variable, by log (VOLUME) and other independent variables: 1979-80

Independent variable	Procedure							
	Transurethral resection of the prostate	Femur fracture reduction	Resection of the intestine	Cholecystectomy	Inguinal hernia repair	Coronary artery bypass	Hip arthroplasty (total)	Hip arthroplasty (other)
	Sign of the regression coefficient (p-value)							
Age	Pos. (<.001)	Pos. (<.001)	Pos. (<.001)	Pos. (<.001)	Pos. (<.001)	Pos. (<.001)	Pos. (<.001)	Pos. (<.001)
Sex ¹	(¹⁰)	Pos. (.350)	Pos. (.004)	Pos. (.001)	Neg. (.053)	Neg. (.190)	Pos. (.246)	Pos. (.554)
Age x sex	(¹⁰)	Pos. (.546)	Neg. (.022)	Neg. (.007)	Pos. (.037)	Pos. (.299)	Neg. (.427)	Pos. (.465)
Emergency ²	(¹⁰)	(¹⁰)	Pos. (<.001)	Pos. (<.001)	Pos. (.008)	(¹⁰)	(¹⁰)	(¹⁰)
Diagnosis ³	Pos. (<.001)	(¹⁰)	¹¹ Neg. (<.001)	(¹⁰)	¹³ Pos. (<.001)	(¹⁰)	¹⁴ Pos. (<.001)	¹⁴ Neg. (.165)
	—	—	¹² Neg. (<.001)	—	—	—	—	—
Geographic region: ⁴								
Northeast	Pos. (.010)	Pos. (<.001)	Pos. (.362)	Pos. (.993)	Neg. (.420)	Pos. (.096)	Neg. (.364)	Pos. (.060)
Southern	Pos. (.014)	Pos. (.006)	Pos. (.039)	Neg. (.775)	Pos. (.905)	Pos. (.337)	Pos. (.658)	Pos. (.008)
North-Central	Pos. (.017)	Pos. (.003)	Pos. (.083)	Neg. (.698)	Pos. (.221)	Pos. (.031)	Neg. (.789)	Pos. (.003)
MED ⁵	Pos. (.390)	Neg. (.926)	Pos. (.015)	Pos. (.601)	Pos. (.002)	Pos. (.685)	Neg. (.037)	Neg. (.664)
SMSA ⁶	Neg. (.954)	Neg. (.102)	Pos. (.286)	Neg. (.247)	Neg. (.762)	(¹⁰)	Pos. (.696)	Pos. (.233)
PROP ⁷	Pos. (.379)	Pos. (.472)	Pos. (.536)	Pos. (.283)	Neg. (.188)	(¹⁰)	Pos. (.360)	Pos. (.773)
BEDS ⁸	Pos. (.984)	Neg. (.434)	Neg. (.315)	Pos. (.252)	Neg. (.455)	Pos. (.193)	Neg. (.807)	Pos. (.811)
Log (VOLUME) ⁹	Neg. (.017)	Neg. (.723)	Neg. (<.001)	Neg. (.431)	Neg. (.184)	Neg. (.031)	Neg. (.759)	Neg. (.043)

¹ Sex: 1 = Male; 0 = Female.

² Emergency admission is defined as one in which surgery was performed on the same day as admission to the hospital.

³ Diagnosis: 1 = Presence of indicated diagnosis; 0 = Absence.

⁴ Geographic region: 0 = Patient not residing in that region. Northeast; 1 = Patient residing in Northeast region. South: 1 = Patient residing in South region. North-Central: 1 = Patient residing in North-Central region.

⁵ MED is medical school affiliation: 1 = Medical school affiliation; 0 = No affiliation.

⁶ SMSA is standard metropolitan statistical area. 1 = Location in SMSA. 0 = Location other than SMSA.

⁷ PROP is proprietary: 1 = Proprietary hospital; 0 = Nonproprietary hospital.

⁸ BEDS refer to bed size.

⁹ VOLUME is inflated by a factor of 5.

¹⁰ Not applicable.

¹¹ 1 = Principal diagnosis of cancer; 0 = Other principal diagnosis.

¹² 1 = Principal diagnosis of diverticula of intestine; 0 = Other principal diagnosis.

¹³ 1 = Principal diagnosis of obstructed hernia; 0 = Other principal diagnosis.

¹⁴ 1 = Principal diagnosis of femur fracture; 0 = Other principal diagnosis.

SOURCE: Health Care Financing Administration, Bureau of Data Management and Strategy: Data from the Medicare Statistical System, 1979-80.

Table 6

Correlations between procedure-specific volume and beds, and procedure-specific volume and general surgical volume, by type of procedure: 1979-80

Procedure	Procedure-specific volume and beds	Procedure-specific volume and general surgical volume
Transurethral resection of the prostate	.62	.79
Femur fracture reduction	.63	.65
Resection of the intestine	.67	.80
Cholecystectomy	.66	.79
Inguinal hernia repair	.63	.76
Coronary artery bypass	.35	.29
Hip arthroplasty (total)	.42	.55
Hip arthroplasty (other)	.60	.72

NOTE: Coefficients of correlation between beds and general surgical volume were between .78 and .86 for all procedures.

SOURCE: Health Care Financing Administration, Bureau of Data Management and Strategy; Data from the Medicare Statistical System, 1979-80.

patients residing in the West. (The West is the region excluded from the regression and therefore each regional dummy variable represents a comparison of that region with the West.) Patients undergoing TURP, hip arthroplasty (other), and femur fracture reduction in the West exhibit significantly better outcomes than patients undergoing the same procedures in most or all of the other three regions ($p < .05$ for the regional dummy variables). Similarly, for resection of the intestine and coronary artery bypass, patients in the West had the lowest mortality, although only one regional coefficient for each procedure is significant at the .05 level. These findings confirm those obtained from earlier univariate analyses on the same data set (Lubitz, Riley, and Newton, 1985). These findings are also similar to those reported by Luft, who found significantly better outcomes in the West for 7 out of 12 procedures studied. No evidence of a regional effect on mortality among cholecystectomy, hip arthroplasty (total) or inguinal hernia repair patients is shown in Table 5.

Medical school affiliation exhibits a positive association with mortality at the .05 level for resection of the intestine and inguinal hernia repair. For hip arthroplasty (total) patients, however, medical school affiliation is significantly associated with favorable outcomes. An often cited explanation for poorer outcomes in hospitals affiliated with a medical school is that the patients undergoing operations in these facilities tend to be in poorer health than those treated elsewhere. Our data do not permit a conclusion on this assertion. The dummy variable designating location of the hospital in an SMSA exhibits no consistent relationship for this set of procedures. Proprietary status yields positive regression coefficients for six of seven procedures, but none of the coefficients achieves statistical significance. Higher mortality is significantly

associated ($p < .001$) with a principal diagnosis of cancer for TURP patients, with obstructed hernia for inguinal hernia repair patients, and with femur fracture for hip arthroplasty (total) patients. Patients undergoing resection of the intestine exhibit more favorable outcomes with a principal diagnosis of diverticula of the intestine, and, surprisingly, with a principal diagnosis of cancer. It is possible that the remaining patients, who exhibited a variety of principal diagnoses and who constituted only 35 percent of all resection patients, suffered from multiple conditions, and that cancer or some other serious illness was present, but not coded as the principal diagnosis associated with the hospitalization. It is also possible that many of the residual patients were nonelective cases, and constituted poorer risks for surgery than the cancer patients, who are commonly elective cases.

The predicted probability of death within 60 days of surgery is given for various levels of VOLUME for resection of the intestine, coronary artery bypass, TURP, and hip arthroplasty (other) (Table 7), which are the four procedures for which VOLUME was significant in Table 5. The probabilities are evaluated at the mean of the other independent variables. The predicted probability of death decreases substantially as VOLUME increases for resection of the intestine, coronary artery bypass, and hip arthroplasty (other). For resection of the intestine, a change in VOLUME from 10 to 25 operations decreases the predicted probability of death from .109 to .097 (a 11.0 percent decrease). An increase in VOLUME from 10 to 40 operations decreases the predicted probability of death from .109 to .091, or 16.5 percent. Similarly, an increase in VOLUME of coronary artery bypass

Table 7

Estimated probability of death within 60 days of surgery for Medicare enrollees 65 years of age or over, by type of procedure and volume, controlling for other independent variables: 1979-80

Procedure and volume	Estimated probability of death
Resection of the intestine	
10	.109
25	.097
40	.091
Coronary artery bypass	
20	.067
40	.061
80	.056
Transurethral resection of the prostate	
20	.027
50	.025
100	.023
Hip arthroplasty (other)	
10	.081
25	.076
50	.072

SOURCE: Health Care Financing Administration, Bureau of Data Management and Strategy; Data from the Medicare Statistical System, 1979-80.

Table 8

Comparison of LOG (VOLUME) effects in logistic regression models using death during surgical stay and death within 60 days of surgery as the dependent variable, by type of procedure: 1979-80

Procedure	Death during surgical stay		Death within 60 days of surgery		Percent of cases dying	
	Coefficient	p-value	Coefficient	p-value	During surgical stay	Within 60 days of surgery
Transurethral resection of prostate	-.218	<.001	-.087	.017	1.0	2.9
Femur fracture reduction	-.064	.187	-.013	.723	5.0	10.2
Resection of intestine	-.176	<.001	-.147	<.001	7.2	11.0
Cholecystectomy	-.099	.072	-.034	.431	2.5	4.2
Inguinal hernia repair	-.191	.117	-.096	.184	0.5	1.5
Coronary artery bypass	-.149	.033	-.131	.031	4.7	6.5
Hip arthroplasty (total)	-.273	.038	-.030	.759	1.1	2.3
Hip arthroplasty (other)	-.163	.003	-.078	.043	4.2	8.9

SOURCE: Health Care Financing Administration, Bureau of Data Management and Strategy: Data from the Medicare Statistical System, 1979-80.

surgeries from 20 to 40 results in a decrease in predicted mortality from .067 to .061 (a 9.0 percent decrease). An increase in VOLUME from 20 to 80 coronary artery bypass operations per year results in a decrease in the predicted probability of death from .067 to .056 (a 16.4 percent decrease). For hip arthroplasty (other), an increase in VOLUME from 10 to 25 reduces the predicted probability of death from .081 to .076 (or 6.2 percent); an increase from 10 to 50 operations reduces it from .081 to .072 (11.1 percent). For TURP, the impact of a change in VOLUME is considerably less. An increase in VOLUME from 20 to 100 operations produces a decrease in predicted probability of death from .027 to .023. Although this represents a 14.8 percent decrease, the decline is not very substantial in absolute terms because of the relatively low mortality rate associated with TURP.

The models are not suitable for predicting individual probabilities of death. In these cases, outcome can be affected by many additional variables that cannot be controlled in a model of this type.

As mentioned earlier, previous studies of the volume-mortality relation looked at deaths during the initial stay for surgery; we developed an additional set of models using in-hospital deaths as the dependent variable. As indicated in Table 8, the association between surgical volume and in-hospital mortality is much stronger than it is between volume and mortality within 60 days of surgery. Five of the eight procedures show a statistically significant relation between volume and in-hospital deaths and the relation approaches significance for a sixth. For TURP, hip arthroplasty (other), and hip arthroplasty (total), the magnitude of the regression coefficients is much greater for the model using in-hospital deaths, and the *p*-values are considerably lower; both facts indicate a stronger volume-mortality relationship for in-hospital mortality than mortality within 60 days of surgery. For coronary artery bypass and resection of the intestine, the two models exhibit a similar association between volume and mortality. This is not too surprising, given the fact that many deaths following these procedures occur in the hospital (Table 8). Luft (1980) found a significant inverse association between in-hospital mortality and surgical

volume for all four of the procedures common to both studies: TURP, cholecystectomy, coronary artery bypass, and hip arthroplasty (total).

The reasons for a stronger relationship between in-hospital mortality and surgical volume are not clear. Because most deaths that occur during the surgical stay happen earlier than 60 days following surgery, it is possible that the effects associated with greater volume are manifested shortly after surgery, and may not be as evident for longer postoperative periods. It is also possible that high-volume hospitals may tend to discharge ill patients sooner than low-volume hospitals, and that more of their postoperative deaths may occur following discharge from the surgical stay. This might happen if, for example, many low-volume hospitals were located in rural areas where there are relatively few opportunities to transfer patients to other hospitals or nursing homes.

Conclusions

Our findings indicate that for the aged, increased institutional volume of surgery is associated with lower postsurgical mortality for some procedures. A significant inverse association between volume and mortality was found for resection of the intestine, TURP, coronary artery bypass, and hip arthroplasty (other). No statistically significant relationship was found between volume and mortality within 60 days of surgery for cholecystectomy, inguinal hernia repair, hip arthroplasty (total), or femur fracture reduction. The inverse association between volume and mortality appears to be much stronger when in-hospital deaths, rather than death within 60 days of surgery, are used as the mortality measure. Although procedure-specific volumes were used in the analysis, it is possible that some of the effects attributed to procedure-specific surgical volume may result from general surgical volume of any kind.

Although every attempt was made to control for appropriate patient characteristics, the results could be attributable to differences in case mix not measured by this study. For example, Flood, Scott, and Ewy (1984a and 1984b) and Luft, Bunker, and Enthoven (1979) found expected mortality to be generally higher in low-volume hospitals. High

mortality in low-volume hospitals could reflect the fact that their caseload may contain a higher proportion of emergency or urgent cases. This might happen if nonemergency cases are often referred to specialty hospitals. Lower mortality in high-volume hospitals could reflect less stringent indications for surgery and hence better results. On the other hand, it is also commonly alleged that "sicker" patients are often referred to larger specialty hospitals and that their case mix is worse than that for less experienced community hospitals. Although we have no way of determining whether this is the case, we can say that, if true, this phenomenon would tend to mask any underlying mortality-volume relationship rather than falsely suggest it.

There are outcomes other than mortality that may be affected by institutional surgical experience. Farber, Kaiser, and Wenzel (1981) found a significant inverse relationship between surgical volume and incidence of postoperative wound infection for six of seven procedures including cholecystectomy, colon resection, and herniorrhaphy. Some of our study procedures, particularly those with low mortality, may be sensitive to surgical volume in terms of morbidity outcomes or patient functioning, but not mortality.

Specific causation cannot be attributed to the volume-mortality relationship found in our study. One possible interpretation is that greater surgical experience leads to improved technique, which results in lower mortality rates over time. Another interpretation is that hospitals that exhibit the best outcomes will attract more referrals and thereby increase their surgical volume. Luft (1980) found evidence of this effect in his study. Determination of causal factors was beyond the scope of our study, however.

Further research is needed on the reasons for the mortality-volume relationship, particularly the importance of the operating surgeon's experience. For example, it is possible that hospitals with high surgical volume have more skilled or experienced surgeons on their medical staffs, which would account for their superior outcomes. They may also have more specialists, as opposed to general practitioners. The role of the anesthesiologist, the operating room team, and other hospital personnel needs to be investigated. High-volume hospitals may recruit more skilled staffs, or, conversely, the personnel at such hospitals may develop superior skills with increased experience. It is also possible that high-volume hospitals may have more specialized facilities and equipment (e.g., ICU's, CCU's). It is also important to determine why some hospitals have low surgical volume. Some of these hospitals may be geographically isolated, and it may not be practical for them to either increase their volume or to refer certain kinds of cases to other hospitals.

The results of this study indicate that hospitals with low surgical volumes should be reluctant to undertake certain procedures on a nonemergency basis. Although our study does not pinpoint the

separate effects of specific surgical volume and general surgical volume, our findings support the concept of establishing guidelines that include minimum volume levels for hospitals that undertake certain kinds of surgeries. Whether minimum levels are developed for specific procedures or operations in general would depend on the nature of each procedure. Many other factors must also be taken into consideration in establishing recommended minimum levels, such as types of facilities and equipment and availability of services elsewhere.

Luft, Bunker, and Enthoven (1979) have called for regionalization of certain procedures, based on their volume-mortality data. Our results provide support for this idea. For example, our study found a significant volume-mortality relationship for coronary artery bypass, resection of the intestine, and hip arthroplasty (other), which are relatively complex procedures with substantial mortality rates among the elderly. It is these types of operations that would be the most likely candidates for regionalization. Before these or other specific procedures are recommended for regionalization, policymakers should do detailed analyses to insure that the volume-mortality relationship is not due to differences in patient risk factors that could not be taken into account in the present study.

The findings also suggest the desirability of making information available to patients and referring physicians on the number of operations by hospital and by surgeon as a guide to the choice of hospital and surgeon. Information on the outcomes of surgery, of course, would be even more valuable. But, at this stage, we are not sure if there is enough understanding of the factors that determine outcome, especially at the physician level, to assure the correct interpretation of outcome data. Without such understanding, publicizing outcome data may create an incentive for physicians to avoid risky cases. A goal of future research should be the development of sufficient understanding of the factors determining outcome to enable dissemination of outcome data and to guide action to correct instances of less than optimal outcome.

These findings also have possible applications under Medicare's new prospective payment system (PPS) for hospitals. There has been concern that under PPS the financial incentives for hospitals to economize might adversely affect quality of care. On the other hand, PPS could encourage hospitals to specialize in the procedures they do most efficiently and thereby increase the concentration of particular procedures in a smaller number of hospitals. Because of the volume-outcome relationship, this could improve surgical outcomes. Under PPS, HCFA is collecting information on up to three procedures and five diagnoses per hospital stay on 100 percent of inpatient stays in short-stay hospitals. Surgical outcomes could be monitored on a hospital-specific or area-specific basis, using HCFA's expanded data system. If poor outcomes following certain procedures are associated with a given hospital, a low

surgical volume for these procedures may, in the context of other information, explain the reason for these unfavorable outcomes. Solutions might be found, in cooperation with other area hospitals, that would guarantee a minimum number of procedures in any hospital receiving Medicare reimbursement for these procedures.

Lastly, more favorable outcomes were experienced by patients in the West for five out of the eight procedures studied. Because death within 60 days of surgery is the mortality measure used in this study, rather than in-hospital deaths, this finding is not a function of the fact that length of stay tends to be shorter in the West. Further research is needed to determine why patients in the West fare so much better following surgery.

Acknowledgments

The authors wish to thank Marian Gornick and Allen Dobson of the Health Care Financing Administration (HCFA), and Dr. Benjamin Barnes of Harvard University for their excellent comments on earlier drafts of the article. We also wish to thank Marilyn Newton of the Bureau of Data Management and Strategy, HCFA, who provided programming support. Lastly, we wish to express our appreciation to Lee Sadler for her administrative assistance.

References

- Commission on Professional and Hospital Activities: Cholecystectomy mortality. *PAS Reporter*, Vol. 8, No. 8. Apr. 20, 1970.
- Farber, B., Kaiser, D., and Wenzel, R.: Relation between surgical volume and incidence of postoperative wound infection. *N Engl J Med* 305 (4):200-204, July 23, 1981.
- Flood, A.B., Scott, W.R., and Ewy, W.: Does practice make perfect? Part I, The relation between hospital volume and outcomes for selected diagnostic categories. *Med Care* 22 (2):98-114, Feb. 1984a.
- Flood, A.B., Scott, W.R., and Ewy, W.: Does practice make perfect? Part II, The relation between volume and outcomes and other hospital characteristics. *Med. Care* 22 (2):115-125, Feb. 1984b.
- Gersh, B., Kronmal, R., Frye, R., et al.: Coronary arteriography and coronary artery bypass surgery, Morbidity and mortality in patients ages 65 years or older—A report from the coronary artery surgery study. *Circulation* 67 (3):483-491, Mar. 1983.
- Harrell, F.: *The LOGIST Procedure*. In *SUGI Supplemental Library User's Guide*, 1983 Edition. SAS Institute, Inc., Cary, North Carolina, 1983.
- Jensen, J., and Tondevold, E.: Mortality after hip fractures. *Acta Orthop. Scand.* 50:161-167, 1979.
- Lubitz, J., and Deacon, R.: The rise in the incidence of hospitalizations for the aged, 1967 to 1979. *Health Care Financing Review*. Vol. 3, No. 3. HCFA Pub. No. 03141. Office of Research, Demonstrations, and Statistics, Health Care Financing Administration. Washington. U.S. Government Printing Office, Mar. 1982.
- Lubitz, J., Riley, G., and Newton, M.: Outcomes of surgery in the Medicare aged population, Mortality after surgery. *Health Care Financing Review*. Vol. 6, No. 4. HCFA Pub. No. 3205. Office of Research and Demonstrations, Health Care Financing Administration. Washington, U.S. Government Printing Office, Summer 1985.
- Luft, H.: The relation between surgical volume and mortality, An exploration of causal factors and alternative models. *Med Care* 18 (9):940-959, Sept. 1980.
- Luft, H., Bunker, J., and Enthoven, A.: Should operations be regionalized? The empirical relation between surgical volume and mortality. *N Engl J Med* 301 (25):1364-1394, Dec. 20, 1979.
- National Academy of Sciences, National Research Council, Subcommittee on the National Halothane Study of the Committee on Anesthesia: Summary of the national halothane study. *JAMA*, 197 (10):121-134, Sept. 5, 1966.
- National Academy of Sciences: *Reliability of Medicare Hospital Discharge Records*. Contract No. SSA 600-7-0159. Prepared for the Department of Health, Education, and Welfare. Washington, D.C. Institute of Medicine, Nov. 1979.
- National Center for Health Statistics, Division of Health Care Statistics: *Detailed Diagnoses and Surgical Procedures for Patients Discharged from Short-Stay Hospitals, United States, 1979*. DHHS Pub. No. (PHS) 82-1274-1. Public Health Service. Washington. U.S. Government Printing Office, Jan. 1982.
- Pokras, R.: Surgical and nonsurgical procedures in short-stay hospitals, United States, 1979. *Vital and Health Statistics*. Series 13, No. 70. DHHS Pub. No. (PHS) 83-1731. Public Health Service. Washington. U.S. Government Printing Office, Feb. 1983.
- Schumacher, D., and Horn, S.: *Hospital Death Rates, Analyses from a Statewide Data System*. Contract No. 600-76-01-40 and Grant 18-P-97031. Technical Report No. 308. Prepared for DHEW by Johns Hopkins University School of Hygiene and Public Health, Baltimore, MD, Dec. 1978.
- Stanford Center for Health Care Research: *Study of Institutional Differences in Postoperative Mortality*. Report No. NCHSR 76-318. Contract No. PH 43-63-65, National Academy of Sciences, Wash. Dec. 1974.