Associations Between Hospital and Surgeon Procedure Volumes and Patient Outcomes After Ovarian Cancer Resection

Deborah Schrag, Craig Earle, Feng Xu, Katherine S. Panageas, K. Robin Yabroff, Robert E. Bristow, Edward L. Trimble, Joan L. Warren

Background: **Strong associations between provider (i.e., hospital or surgeon) procedure volumes and patient outcomes have been demonstrated for many types of cancer operation. We performed a population-based cohort study to examine these associations for ovarian cancer resections.** ***Methods:*** **We used the Surveillance, Epidemiology, and End Results (SEER)–Medicare linked database to identify 2952 patients aged 65 years or older who had surgery for a primary ovarian cancer diagnosed from 1992 through 1999. Hospital- and surgeon-specific procedure volumes were ascertained based on the number of claims submitted during the 8-year study period. Primary outcome measures were mortality at 60 days and 2 years after surgery, and overall survival. Length of hospital stay was also examined. Patient age at diagnosis, race, marital status, comorbid illness, cancer stage, and median income and population density in the area of residence were used to adjust for differences in case mix. All *P* values are two-sided.** ***Results:*** **Neither hospital- nor surgeon-specific procedure volume was statistically significantly associated with 60-day mortality following primary ovarian cancer resection. However, differences by hospital volume were seen with 2-year mortality; patients treated at the low-, intermediate-, and high-volume hospitals had 2-year mortality rates of 45.2% (95% confidence interval [CI] = 42.1% to 48.4%), 41.1% (95% CI = 38.1% to 44.3%), and 40.4% (95% CI = 37.4% to 43.4%), respectively. The inverse association between hospital procedure volume and 2-year mortality was statistically significant both before (*P* = .011) and after (*P* = .006) case-mix adjustment but not after adjustment for surgeon volume. Two-year mortality for patients treated by low-, intermediate-, and high-volume surgeons was 43.2% (95% CI = 40.7% to 45.8%), 42.9% (95% CI = 39.5% to 46.4%), and 39.5% (95% CI = 36.0% to 43.2%), respectively; there was no association between 2-year mortality and surgeon procedure volume, with or without case-mix adjustment. After case-mix adjustment, neither hospital volume (*P* = .031) nor surgeon volume (*P* = .062) was strongly associated with overall survival.** ***Conclusion:*** **Hospital- and surgeon-specific procedure volumes are not strong predictors of survival outcomes following surgery for ovarian cancer among women aged 65 years or older. [J Natl Cancer Inst 2006;98:163–71]**

Compelling evidence from multiple studies *(1–5)* suggests that cancer patients whose surgical resections are performed in hospitals with large case volumes have better outcomes than patients treated in hospitals with low case volumes. Results of another study *(6)* suggest that the case volume of the individual surgeon is a more powerful determinant of outcome than the case volume of hospital in which a resection is performed. The sustained interest in volume–outcome studies in the medical literature and in the lay press has led patients to ask practical questions about how they should weigh information about the case volumes of the hospital and of the individual surgeon in their decisions about where and from whom to seek medical care. From a policy perspective, the most important sources of variation in outcomes need to be identified to optimize the quality of care.

We examined the roles of surgeon procedure volume and hospital procedure volume as determinants of outcomes for a population-based cohort of female Medicare beneficiaries diagnosed with epithelial ovarian cancer during the 1990s. We sought to determine whether patients treated by high-volume providers (i.e., hospitals and surgeons) achieve better outcomes than patients treated by low-volume providers. In particular, we wanted to know whether selecting an ovarian cancer surgeon on the basis of his or her case volume represents a good strategy for optimizing outcomes, independent of the surgeon's specialty (i.e., gynecologic oncology, general gynecology, or general surgery). We anticipated that there might be some general gynecologists and general surgeons who achieved outcomes comparable to those of gynecologic oncologists because they performed ovarian cancer operations frequently.

Methods

Data Sources

We identified a cohort of elderly ovarian cancer patients from a database *(7)* that links the Surveillance, Epidemiology, and End Results (SEER) population-based cancer registries with a Center for Medicare Services's health care claims. We used the unique provider identification numbers (UPINs) included on all

Affiliations of authors: Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, NY (DS, FX, KSP); Center for Outcomes and Policy Research, Dana-Farber Cancer Institute, Boston, MA (CE); Department of Gynecology and Obstetrics, The Johns Hopkins Medical Institutions, Baltimore, MD (REB); Applied Research Program (KRY, JLW), Cancer Therapy and Evaluation Program (ELT), National Cancer Institute, Bethesda, MD.

Correspondence to: Deborah Schrag, MD, Department of Epidemiology and Biostatistics, Box 221, Memorial Sloan-Kettering Cancer Center, 1275 York Ave., New York, NY 10021 (e-mail: schragd@mskcc.org).

See "Notes" following "References."

DOI: 10.1093/jnci/djj018

claims submitted for reimbursement to Medicare since 1991 to obtain information about the patients' providers (i.e., hospitals and surgeons).

The SEER registries ascertain all incident cancer cases diagnosed in five states and six U.S. metropolitan areas, which together represent approximately 14% of the U.S. population *(8)*. The SEER program collects information on each incident cancer, including the primary site and histology [classified according to the International Classification of Disease for Oncology, 2nd edition *(9)*], the tumor stage at diagnosis, and patient demographics *(8,9)*. SEER data do not include detailed information about cancer treatment; however, this information can be ascertained for Medicare beneficiaries from billing claims. Given that ovarian cancer typically strikes women in their 60s and 70s *(11)* and that the Medicare program provides health insurance for 97% of the U.S. population aged 65 years or older *(12)*, linkage of the SEER registries with Medicare claims data allowed us to identify a large, nationally representative cohort of women with this disease *(7)*. Among SEER registry patients aged 65 years or older, 94% have been linked to their Medicare records *(12)*. Data from the 2000 U.S. Census have also been linked to SEER–Medicare data, allowing us to characterize patients represented in the SEER–Medicare database with respect to sociodemographic factors.

We used Medicare Provider Analysis and Review (MEDPAR) files to obtain details of all hospitalizations for persons eligible for Medicare Part A. To receive payment, hospitals submit claims to Medicare that code up to 10 diagnoses and 10 procedures that are classified according to the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) *(13)*. For the 96% of Medicare beneficiaries who opt for Part B coverage, claims for care delivered in hospital outpatient departments or in physicians' offices are also recorded. Medicare documents the date of death for its beneficiaries using information provided by the Social Security Administration.

Cohort Definition

All Medicare-enrolled patients aged 65 years or older who were diagnosed with primary ovarian cancer while residing in a area covered by the SEER program between 1992 and 1999 were potentially eligible for inclusion in our study. We restricted our cohort to patients with a histologic diagnosis consistent with epithelial cell tumors and excluded patients with germ-cell tumors and ovarian tumors of borderline malignancy. We also excluded patients whose diagnoses were noted exclusively on death certificates or at autopsy as well as those for whom the month of diagnosis was not known. We excluded the 24% of ovarian cancer patients living in areas covered by SEER programs who were enrolled in a health maintenance organization (HMO) at diagnosis because HMOs do not report claims detailing the specific procedures and noncancer diagnoses to the Center for Medicare Services. All patients in our cohort were enrolled continuously in Medicare Parts A and B after diagnosis. Our sample size is slightly smaller than that of the companion paper by Earle et al. *(14)* because our analysis required that both the surgeon and hospital be identified for each patient.

Identification of Ovarian Cancer Operations

Both SEER registries and Medicare record information about a patient's initial surgery. In collaboration with a panel of four gynecologic surgeons (Diane C. Bodurka, Department of Gynecologic Oncology, University of Texas M. D. Anderson Cancer Center; Michael Carney, Department of Obstetrics, Gynecology, and Women's Health, Kapiolani Medical Center for Women and Children; ELT; and REB), we identified the SEER variables and the Medicare billing codes that correspond to a primary surgery for ovarian cancer and searched for those variables and diagnosis and procedure codes in 1) inpatient claims submitted by hospitals using ICD-9-CM *(13)*; 2) claims submitted by surgeons using the Current Procedural Terminology (CPT) codes *(15,16)*; and 3) information from the SEER registries about the type of cancer-directed surgery performed *(8)*. Our analysis included only patients who underwent surgery at hospitals located in one of the nine states covered by one of the 11 SEER registries because we could not reliably measure procedure volume at institutions outside SEER areas.

Specification of Cohort Characteristics

Demographic variables, including age at diagnosis and race, were defined on the basis of information ascertained by the SEER registrars. Cancer stage was assigned according to the American Joint Committee on Cancer (AJCC) classification schema recorded by SEER *(17)*. We used the Romano modification *(18)* of the Charlson comorbidity index *(10)* to ascertain comorbidities among our cohort. To categorize comorbidities, we examined all inpatient Medicare claims for the 12 months prior to the index surgical admission as well as claims filed during the index admission and used these records to assign to each patient the maximal comorbidity observed. Information about marital status, median income in the census tract of patient residence, and the population density in the area of residence was obtained from Medicare files.

Outcomes After Ovarian Cancer Surgery

Outcome measures included 60-day and 2-year postoperative mortality and overall survival. Survival was defined as the interval from the date of hospitalization for resection until the date of death as reported to Medicare or December 31, 2001 (i.e., when censoring occurred), whichever occurred first. Other outcome measures were the length of hospital stay and the rate of operations that were accompanied by creation of an intestinal stoma (i.e., ostomy rate). Ovarian cancer resections that required an ileostomy or colostomy were identified in SEER–Medicare data on the basis of accompanying ICD-9 or CPT codes. We did not consider other complications as outcomes because we were concerned that their coding might vary according to local physician and hospital practices.

Procedure Volume

Hospitals were ranked by volume according to the total number of ovarian cancer operations performed from January 1, 1992, through December 31, 1999, as identified from the Medicare files. We used an analogous approach to ascertain surgeon-specific procedure volume. Surgeons were ranked according to their total volume of claims for ovarian cancer resections performed on cohort members from January 1, 1992, through December 31, 1999. Patients whose claims lacked a UPIN for the primary surgeon were categorized as having been treated by a surgeon with an unknown procedure volume. We could not determine the

surgeon procedure volume for 452 (13.3%) of the 3404 patients in our original cohort either because there was a hospital claim but no physician claim for the ovarian cancer operation or because the physician claim lacked a UPIN. Therefore, all primary analyses were performed for the subset of 2952 patients for whom both surgeon- and hospital-specific procedure volumes were available. Both the baseline characteristics and outcomes of patients for whom the case volume of the surgeon could not be specified were similar to those for the entire cohort (data not shown).

Postoperative Chemotherapy

We included postoperative chemotherapy initiated within 6 months of the index surgery in our analysis because it is an important component of disease management for women with ovarian cancer, particularly those with advanced disease. We reviewed all Medicare claims for chemotherapy drugs or administration, irrespective of whether administration was done on an inpatient or outpatient basis or on the route of administration (i.e., intravenous or intraperitoneal). Chemotherapy was identified by ICD-9 procedure code 99.25 and CPT codes J9000–J9999, 96400–96549, Q0083, Q0084, and Q0085.

Surgeon Specialty

Information about physician characteristics, particularly the specialty type of the surgeon who performed a cancer-directed procedure, was obtained by linking UPINs in Medicare claims to data in the American Medical Association (AMA) Masterfile. This linkage was performed to determine the extent to which physician specialty type (i.e., gynecologic oncologist, general gynecologist, or general/other surgeon) was related to procedure volume and whether physician specialty might mediate the association between procedure volume and outcomes. A detailed analysis of the association between surgeon specialty and ovarian cancer outcomes using this dataset is reported in the companion article by Earle et al. *(14)*.

Statistical Analysis

For each outcome of interest, we used generalized estimating equation models to perform analyses with and without adjustment for differences in case mix as well as with adjustment for case-mix and surgeon volume (when the primary variable was hospital volume) or with adjustment for case-mix and hospital volume (when the primary variable was surgeon volume). Cox proportional hazards models were used to evaluate the association between volume and overall survival, and the validity of the proportional hazards assumption was confirmed by visual inspection of the volume category- specific Kaplan–Meier survival curves. (Details of the statistical models are available at: http://jncicancerspectrum.oxfordjournals.org/jnci/content/vol98/issue3.) Patient marital status, median income in the census tract of patient residence, and the population density in the area of residence have been associated with access to care *(1,2,4,5)* and thus were included in the multivariable analyses. Case-mix variables—including age at diagnosis, race, AJCC stage at diagnosis, comorbidity, socioeconomic status, and population density in the patient's area of residence—that were found to be associated with outcomes or hospital or surgeon procedure volume in univariate analyses were retained in the final multivariable models. We did not include the reporting SEER registry in the multivariable models because it correlated well *(5)* with population density in the patients' area of residence. We did not include the administration of postoperative chemotherapy in the multivariable models in our analyses of the associations between procedure volume and outcomes of ovarian cancer surgery because such treatment is not an essential attribute of the patient as is stage at diagnosis or comorbidity. However, we did incorporate a variable for postoperative chemotherapy administration into a subsequent analysis to examine whether this intervention might mediate any associations between volume and outcomes.

All statistical analyses were performed using procedure volume as a continuous measure. Also, to facilitate display of our results and to obtain relative risk (RR) estimates from the Cox proportional hazards models, we defined categories of hospital volume and of surgeon-specific procedure volume based on the number of operations performed on members of our cohort during the study period. Because analyses were conducted by considering volume as a continuous measure, the cutpoints for the volume categories did not influence statistical interpretation of results. The cutpoints for the low-, intermediate-, and high-volume categories were selected as follows: We first defined tertiles based on the number of total patients in the cohort and then required all patients operated on by a provider with a specified procedure volume to be assigned the same category. For hospital procedure volume (number of cases during the 8-year study period), these categories were as follows: low (1–12), intermediate (13–28), and high (29–93). Surgeon procedure volume categories were as follows: low (1–3), intermediate (4–19), and high (20–61).

All statistical tests of association were corrected for clustering of outcomes within a particular hospital or a particular surgeon's panel of patients by using a previously described method *(19)*. For binary outcomes, a generalized estimating equation model was fitted to the data to correct for within-hospital or within-surgeon correlation *(20)*. To evaluate the association between overall survival and volume, we constructed a Cox proportional hazards model corrected for either within-surgeon correlation (when hospital volume was the primary predictor) or within-hospital correlation (when surgeon volume was the primary predictor).

To further assess the extent to which outcomes varied among providers, we measured outcomes for the 1015 patients treated at the 27 hospitals in the high-volume category and for the 711 patients treated by the 27 surgeons in the high-volume category. We restricted this analysis to patients who were treated by high-volume providers because analysis of provider-specific variation in outcomes is increasingly unreliable as the number of patients in a particular provider's patient panel decreases. We constructed a test for extrabinomial variation that compared the observed and expected frequencies for mortality outcomes for each high-volume surgeon and hospital. This test identified the presence of additional hospital-to-hospital (or surgeon-to-surgeon) variation that could not be explained by the other factors in the multivariable model. The number of standardized deviations (z scores) of the observed frequency from the expected frequency for each hospital (and each surgeon) was then used to calculate correlation coefficients among the endpoints and to characterize the degree to which an individual surgeon was associated with exceptionally high or low risks for adverse outcomes. These

results are displayed using histograms to juxtapose the observed distribution of outcomes and the distribution expected if outcomes were not associated with volume.

All *P* values are two-sided and were calculated using the Mantel–Haenszel chi-square test. *P* values less than .05 were considered statistically significant. SAS software (version 9.0; SAS Institute Cary, NC) was used for all analyses.

Results

Cohort Characteristics

Using SEER–Medicare data, we identified 2952 patients who had a primary ovarian cancer resection during the 8-year study period and for whom both hospital- and surgeon-specific procedure volumes could be defined. Patient characteristics according to the procedure volume of the hospital where the primary resection was performed and of the surgeon who performed the resection are shown in Table 1. Patients in the three strata of hospital volume were similar with respect to marital status and comorbidity level. However, patients in the three strata of hospital volume differed with respect to age at diagnosis, race, stage at diagnosis, and median income in the census tract of residence. For example, patients who were younger at diagnosis were more likely than patients who were older at diagnosis to be operated on at high-volume institutions. Patients who resided in more affluent areas were more likely than patients who resided in less affluent areas to undergo surgery at high-volume hospitals ($P<.001$), as were patients who were diagnosed with advanced-stage cancer compared with patients who were diagnosed with early-stage cancer ($P = .024$). Residents of rural areas were more likely to have surgery at a low-volume hospital than residents of more urban communities. Each SEER region had patients represented in each volume category. For example, patients who were diagnosed in Iowa (a rural state) made up 14.8% of the cohort but contributed 19.7% of the patients treated at a low-volume and only 12.3% of those treated at a high-volume hospital.

Associations between patient characteristics and surgeon volume followed a pattern similar to that seen for hospital volume. High-volume surgeons were less likely than low-volume surgeons to operate on elderly patients ($P = .001$) or on patients with early-stage disease ($P<.001$). Patients who resided in a low-income census tract ($P = .001$) or in a rural area ($P<.001$) were also less likely to have a high-volume surgeon than those who resided in a high-income census tract or in an urban area. There was no association between surgeon case volume and either patient marital status or comorbidity (Table 1).

Patients treated by high-volume hospitals or surgeons were more likely to receive postoperative chemotherapy than those treated by low-volume hospitals or surgeons ($P<.001$ for both) (Table 1). For example, 62.9% of patients treated at low-volume hospitals received chemotherapy compared with 75.8% of those treated at high-volume hospitals, and 63.6% of patients treated by low-volume surgeons received postoperative chemotherapy compared with 75.7% of those treated by high-volume surgeons.

Primary Outcomes

The primary outcomes of our study are shown in Table 2, and Kaplan–Meier survival curves are shown in Fig. 1. Cohort members had their resections performed at 423 different institutions by 1365 different primary surgeons. Many hospitals (78.0% of total) and surgeons (87.6% of total) had low ovarian cancer–specific procedure volumes, whereas few hospitals (6.4% of total) and surgeons (2.0% of total) had high procedure volumes.

Sixty-day postoperative mortality. Patients operated on at high-volume hospitals had a 2% lower 60-day mortality rate than those operated on at low-volume hospitals (6.6% [95% confidence interval [CI] = 5.2% to 8.3%] versus 8.6% [95% CI = 7.0% to 10.5%]), and those operated on by high-volume surgeons had a 2.7% lower 60-day mortality rate than those operated on by low-volume surgeons (6.5% [95% CI = 4.9% to 8.5%] versus 9.2% [95% CI = 7.8% to 10.8%]). However, there was no evidence of a statistically significant association between either hospital or surgeon procedure volume and 60-day mortality.

Two-year mortality. At 2 years after ovarian cancer resection, patients treated at the low-, intermediate-, and high-volume hospitals had mortality rates of 45.2% (95% CI = 42.1% to 48.4%), 41.1% (95% CI = 38.1% to 44.3%), and 40.4% (95% CI = 37.4% to 43.4%), respectively. The association between hospital procedure volume and 2-year mortality was statistically significant after adjustment for case mix ($P = .006$) but not after adjustment for both case mix and surgeon volume ($P = .088$). Two-year mortality for patients treated by low-, intermediate-, and high-volume surgeons was 43.2% (95% CI = 40.7% to 45.8%), 42.9% (95% CI = 39.5% to 46.4%), and 39.5% (95% CI = 36.0% to 43.2%), and there was no statistically significant association between this outcome and surgeon procedure volume, with or without case-mix adjustment (Table 2). Outcomes had a much stronger association with age and cancer stage at diagnosis. For example, 2-year mortality was 35% (95% CI = 32.1% to 37.6%) among 65–69-year-olds versus 47% (95% CI = 42.9% to 49.2%) among 75–79-year-olds ($P<.001$).

Length of hospital stay and ostomy rates. Patients treated at low-volume hospitals had hospital stays that were 2 days longer than those of patients treated at high-volume hospitals (5.5 days [95% CI = 5.0 days to 6.0 days] versus 3.5 days [95% CI = 3.2 days to 3.9 days]). Patients operated on by low-volume surgeons had hospital stays that were 2.8 days longer than those of patients operated on by high-volume surgeons (6.1 days [95% CI = 5.7 days to 6.5 days] versus 3.3 days [95% CI = 2.9 days to 3.8 days]) (Table 2). The associations between length of stay and procedure volumes were statistically significant both before and after adjustment for case mix ($P<.001$ for each). The association between hospital volume and length of stay remained statistically significant after additional adjustment for surgeon volume, as did the association between surgeon volume and length of hospital stay after additional adjustment for hospital volume. Ostomy rates did not vary with surgeon volume but were associated with hospital procedure volume ($P = .005$) in the adjusted analysis.

Overall survival. We found a modest association between hospital volume and overall survival in a Cox model adjusted for case mix ($P = .031$). When surgeon volume was added to the Cox model, the association was no longer evident ($P = .145$). There was no statistically significant association between surgeon volume and overall survival in the case-mix–adjusted analysis ($P = .062$) or when hospital volume was included in the model ($P = .326$).

When we included administration of postoperative chemotherapy in case-mix–adjusted Cox models, the associations between provider procedure volume and patient outcomes were

Table 1. Characteristics of 2952 ovarian cancer patients according to the procedure volume of the hospital where the primary resection was performed and of the surgeon who performed the resection*

Characteristics	N (%)	Hospital procedure volume†, %				Surgeon procedure volume†, %			
		Low (N = 977)	Intermediate (N = 960)	High (N = 1015)	P‡	Low (N = 1455)	Intermediate (N = 786)	High (N = 711)	P‡
Age at diagnosis, y									
65–69	732 (24.8)	23.2	24.1	27.0	<.001§	23.8	24.9	26.7	.001§
70–74	850 (28.9)	27.5	27.7	31.0		28.2	27.5	31.5	
75–79	767 (25.9)	26.0	25.8	26.1		25.4	26.7	26.4	
≥80	603 (20.4)	23.3	22.4	15.9		22.7	20.9	15.3	
Race									
White	2671 (90.5)	89.1	92.7	89.8	.035	91.7	89.7	88.9	.039
Black	137 (4.6)	4.8	4.0	5.1		3.7	6.2	4.8	
Other	144 (4.9)	6.1	3.3	5.1		4.6	4.1	6.3	
Median income									
First quartile	753 (25.5)	32.0	22.3	22.3	<.001	27.5	29.0	17.6	<.001§
Second quartile	736 (24.9)	25.1	24.9	24.8		25.9	24.1	23.9	
Third quartile	739 (25.0)	23.8	25.2	26.1		24.3	24.4	27.3	
Fourth quartile	724 (24.5)	19.1	27.6	26.8		22.3	22.5	31.2	
Marital status									
Not married‖	1600 (54.2)	56.7	52.3	53.6	.134	54.7	53.2	54.3	.761
Married	1352 (45.8)	43.3	47.7	46.4		45.3	46.8	45.7	
Area of residence¶									
Large metropolitan	1818 (61.6)	55.4	63.2	66.0	<.001	58.4	59.7	70.3	<.001
Metropolitan	652 (22.1)	20.0	25.2	21.2		21.4	21.0	24.6	
Urban	201 (6.8)	9.9	5.3	5.2		8.7	7.0	2.8	
Less urban	225 (7.6)	12.4	4.9	5.6		9.6	9.4	1.7	
Rural	56 (1.9)	2.3	1.4	2.0		2.0	2.9	0.6	
SEER registry									
San Francisco	213 (7.2)	8.2	11.4	2.4	<.001	8.9	6.9	4.1	<.001
Connecticut	371 (12.6)	11.2	8.5	17.7		11.4	7.6	20.4	
Detroit	487 (16.5)	11.0	18.4	20.0		12.7	17.8	22.9	
Hawaii	47 (1.6)	1.5	0.0	3.2		1.2	0.0	4.2	
Iowa	438 (14.8)	19.7	12.5	12.3		17.8	19.5	3.7	
New Mexico	109 (3.7)	5.6	2.0	3.5		3.6	0.6	7.2	
Seattle	346 (11.7)	8.5	14.5	12.2		10.7	12.2	13.2	
Utah	154 (5.2)	4.8	7.4	3.6		6.2	8.0	0.1	
Atlanta	184 (6.2)	5.8	1.5	11.1		4.5	5.0	11.1	
San Jose	130 (4.4)	3.2	6.3	3.8		4.5	6.4	2.0	
Los Angeles	473 (16.0)	20.5	17.6	10.3		18.4	16.0	11.1	
AJCC stage									
I	462 (15.7)	16.4	17.0	13.7	.024	18.3	14.1	12.0	<.001
II	259 (8.8)	10.1	9.1	7.2		10.0	7.8	7.5	
III	1276 (43.2)	43.2	40.3	46.0		39.2	45.8	48.7	
IV	955 (32.3)	30.3	33.6	33.1		32.6	32.3	31.9	
Comorbidity score#									
0	2195 (74.4)	73.2	74.5	75.4	.102	74.0	75.3	74.0	.755
1	532 (18.0)	17.2	19.6	17.3		18.1	17.1	18.9	
2	225 (7.6)	9.6	5.9	7.3		7.8	7.6	7.2	
Postoperative chemotherapy									
No	927 (31.4)	37.1	33.1	24.2	<.001	36.4	28.5	24.3	<.001
Yes	2025 (68.6)	62.9	66.9	75.8		63.6	71.5	75.7	

*Percentages for some categories do not total 100% because of rounding. SEER = Surveillance, Epidemiology, and End Results ; AJCC = American Joint Committee on Cancer.

†Hospital procedure volume categories (number of cases/8-year follow-up): low (1–12), intermediate (13–28), and high (29–93). Surgeon procedure volume categories (number of cases/8-year study observation period): low (1–3), intermediate (4–19), and high (20–61).

‡All *P* values are unadjusted, two-sided, and were calculated using the Mantel–Haenszel chi-square test.

§P_{trend}.

‖Includes patients who were single, widowed, divorced, or separated.

¶Categories are defined by U.S. Census bureau based on residential population in the county of residence. Large metropolitan: >500 000 residents; metropolitan: >250 000 to ≤500 000 residents; urban: >20 000 to ≤250 000 residents; less urban: >2500 to ≤20 000 residents; rural: ≤2500 residents.

#As defined in Romano et al. *(18)*.

Table 2. Outcomes for ovarian cancer patients following surgical resection according to hospital and surgeon procedure volumes*

Outcome	Hospital procedure volume				Surgeon procedure volume			
	Low n = 330 (78.0%)	Intermediate n = 66 (15.6%)	High n = 27 (6.4%)	P†	Low n = 1196 (87.6%)	Intermediate n = 142 (10.4%)	High n = 27 (2.0%)	P†
60-day mortality rate, % (95% CI)	8.6 (6.9 to 9.2)	9.6 (8.0, 10.1)	6.6 (5.9 to 8.0)	.071 .343 .523	9.2 (7.2 to 9.9)	8.0 (7.3 to 9.4)	6.5 (5.2 to 8.0)	.472 .592 .818
2-year mortality rate, % (95% CI)	45.2 (43.2 to 46.9)	41.1 (39.7 to 43.6)	40.4 (38.5 to 42.3)	.011 .006 .088	43.2 (41.6 to 46.4)	42.9 (40.4 to 45.3)	39.5 (38.2 to 41.4)	.228 .087 .330
Ostomy rate, % (95% CI)	17.7 (15.6 to 18.9)	14.8 (13.0 to 15.8)	14.8 (13.6 to 16.1)	.021 .005 .001	16.2 (15.4 to 18.4)	15.1 (14.2 to 16.7)	15.5 (13.5 to 16.0)	.808 .895 .345
Length of hospital stay, days (95% CI)	5.5 (4.2 to 9.2)	5.7 (4.0 to 10.1)	3.5 (3.0 to 8.4)	<.001 <.001 .002	6.1 (4.2 to 10.1)	4.1 (3.8 to 8.6)	3.3 (2.8 to 7.6)	<.001 <.001 <.001

*Hospital procedure volume categories (number of cases per 8-year follow-up): low (1–12), intermediate (13–28), and high (29–93). Surgeon procedure volume categories (number of cases per 8-year study observation period): low (1–3), intermediate (4–19), and high (20–61). CI = confidence interval.

†The top *P* value represents the association between volume and the outcomes without adjustment for other characteristics. The middle *P* value is adjusted for differences in case mix according to the categories shown in Table 1. The bottom *P* value is adjusted for both case mix and surgeon volume (for analyses of hospital volume) and hospital volume (for analyses of surgeon volume). All *P* values were calculated using generalized estimating equations and are two-sided.

further attenuated. For example, the *P* value for the association between hospital volume and overall survival increased from .031 to .362 after adjusting for variation in patterns of chemotherapy use. The *P* value for the association between surgeon volume and survival increased from .062 to .158 after adjusting for postoperative chemotherapy administration.

To illustrate the magnitude of the association between provider procedure volume and patient outcomes after ovarian cancer surgery, we plotted survival curves by category of hospital procedure volume (Fig. 1, A) and by category of surgeon procedure volume (Fig. 1, B). Because cancer stage at diagnosis is such an important prognostic factor, we also stratified the survival curves by whether patients had early (stage I or II) versus late (stage III or IV) stage ovarian cancer at diagnosis. Among patients with stage I or II disease at diagnosis, those operated on at low-volume hospitals had worse survival than those treated at high- or intermediate-volume hospitals (adjusted risk ratio [RR] for women operated on at a low-volume versus high-volume hospitals = 1.37, 95% CI = 0.97 to 1.91; *P* = .07). This association did not change when chemotherapy was included as a covariate in the Cox model (RR = 1.37, 95% CI = 0.98 to 1.92; *P* = .07). Among patients with stage III or IV disease at diagnosis, the survival curves rank by volume category until 28 months, when the survival curve for patients treated at intermediate-volume hospitals became comparable to that for patients treated at high-volume hospitals. The volume–outcome association was somewhat more evident for surgeon procedure volume, particularly for patients with stage III or IV disease, among whom the adjusted risk ratio for those operated on by a low-volume surgeon compared with those operated on by a high-volume surgeon was 1.15 (95% CI = 1.02 to 1.30; *P* = .03). However, this volume–outcome association was attenuated when chemotherapy was included in the model (RR = 1.10, 95% CI = 0.97 to 1.24; *P* = .13).

Two-Year Mortality Rates Among High-Volume Providers

To more comprehensively assess the magnitude of differences among providers, we examined the case mix–adjusted 2-year mortality rates for the 27 hospitals and 27 surgeons in the highest volume categories (Fig. 2). There was no statistically significant difference between the distributions of the observed and the expected 2-year mortality rates among the high-volume hospitals (*P* = .95) and among the high-volume surgeons (*P* = .79) based on extrabinomial variation. This finding suggests that variation in 2-year mortality rates among patients operated on in high-volume

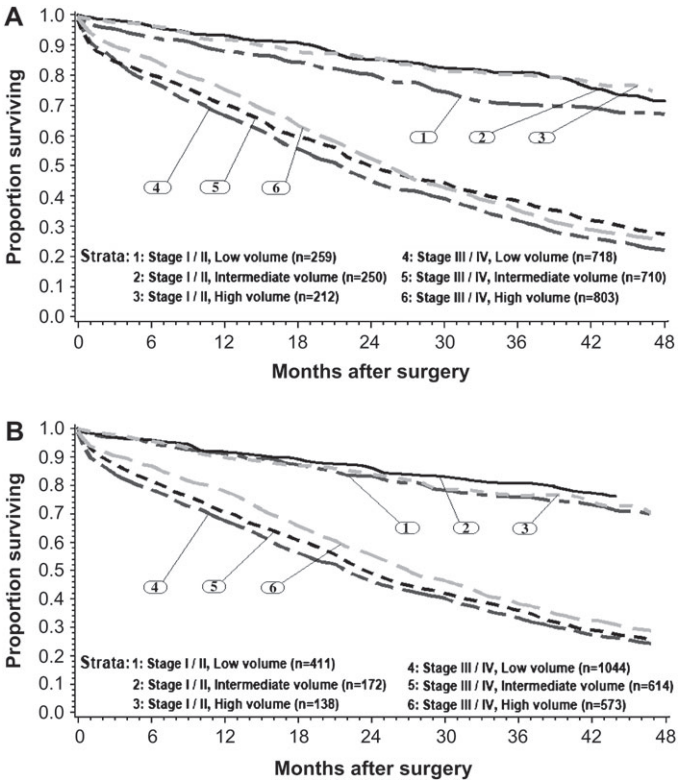

Fig. 1. Overall survival by cancer stage at diagnosis for patients who had surgery for ovarian cancer. **A**) Overall survival according to hospital procedure volume. **B**) Overall survival according to surgeon procedure volume. Cancer stage is according to American Joint Committee on Cancer classification. Hospital procedure volume categories (number of case during the 8-year study period): low (1–12), intermediate (13–28), and high (29–93). Surgeon procedure volume categories (number of case during the 8-year study period): low (1–3), intermediate (4–19), and high (20–61). Numbers on each curve indicate procedure-volume strata.

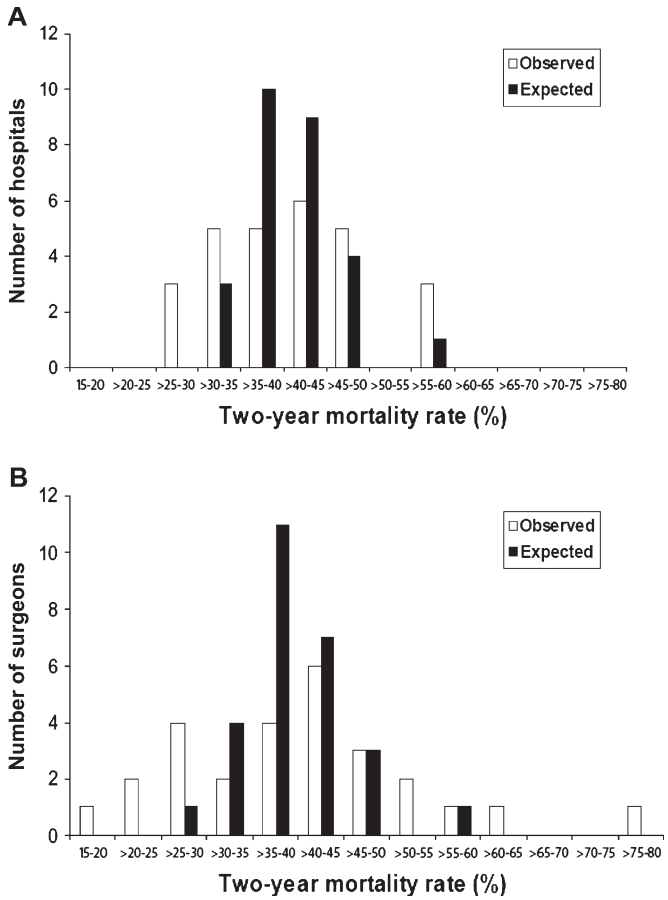

Fig. 2. Observed versus expected 2-year mortality rates for high-volume hospitals and surgeons. **A)** Two-year mortality rates among ovarian cancer patients treated at high-volume hospitals. **B)** Two-year mortality rates among ovarian cancer patients treated by high-volume surgeons. The histograms juxtapose the observed (**white bars**) and expected (**black bars**) number of each hospital's (or each surgeon's) panel of patients who died within 2 years after surgery.

hospitals and/or by high-volume surgeons was no different from that expected based on chance alone and that our findings could not be explained on the basis of clusters of providers with especially stellar or poor performance. Similar findings were obtained for 60-day mortality and ostomy rates (data not shown).

Provider Specialty and Procedure Volume

Ovarian cancer surgery may be performed by gynecologic oncologists, gynecologists, or general surgeons. Whereas gynecologic oncologists represented only 7% of the surgeons in our sample, they operated on 28% of the patients in the cohort and therefore had a substantially higher mean case volume (11 cases/surgeon) than either gynecologists (two cases/surgeon), who represented 55% of the surgeons and operated on 51% of the cohort, or general surgeons (two cases/surgeon), who represented 38% of the surgeons and operated on 21% of the cohort. Of the 83 gynecologic oncologists in our sample, 17 were in the high-volume category (63% of the 27 high-volume surgeons), 37 were in the intermediate-volume category (26% of the 142 intermediate-volume surgeons), and 29 were in the low-volume category (2% of the 1196 low-volume surgeons). This finding emphasizes the relationship between surgeon specialty and case volume. Similarly, of the 83 gynecologic oncologists, 30 worked in one of 212 high-volume hospitals, 17 worked in one of 495 intermediate-volume hospitals, and 36 worked in one of 658 low-volume hospitals.

Discussion

Volume–outcome studies have been used to quantify the degree of variation in outcomes, thereby focusing scrutiny on the precise aspects of care that lead to decreased morbidity and mortality. We found that neither hospital- nor surgeon-specific procedure volume was strongly associated with outcomes among a population-based cohort of U.S. Medicare patients with primary ovarian cancer. Although patients operated on at high-volume hospitals had a 2% lower 60-day mortality rate than those operated on at low-volume hospitals, and those operated on by high-volume surgeons had a 2.7% lower 60-day mortality than those operated on by low-volume surgeons, these differences were not statistically significant after adjusting for the measurable differences in the case mix. By contrast, we found a statistically significant association between 2-year mortality and hospital procedure volume that persisted after adjustment for differences in the case mix. The association between surgeon volume and 2-year mortality was not statistically significant.

Our analyses suggest that, for ovarian cancer, differences in survival outcomes within the Medicare population are largely attributable to baseline differences in the characteristics of patients treated by low- and high-volume providers (either individual surgeons or hospitals). Specifically, the factors most strongly associated with poor outcome were patient age at diagnosis and advanced clinical stage. The absolute magnitudes of the hospital volume–outcome differences we observed were modest compared with those observed for higher-risk operations to treat pancreatic, esophageal, or lung cancers *(1,2)*. For ovarian cancer, prior studies from Scandinavian countries that have national cancer registries *(24,25)* and from Japan *(26)* have reported that patients who had their ovarian cancer surgery at high-volume hospitals had lower postoperative mortality than those treated at low-volume hospitals.

We found that most surgical treatment for this cohort of elderly women with ovarian cancer was delivered by surgeons who performed such operations only occasionally. This finding is consistent with results reported by Bristow et al. *(27)*, who used Maryland hospital discharge data to assess surgeon volume. Practice guidelines and professional societies (e.g., the Society of Gynecologic Oncologists) encourage women with ovarian cancer to obtain care from surgeons who specialize in this disease in part because of the perception that subspecialty surgeons' focus on gynecologic surgery and higher case volumes enable them to achieve superior surgical outcomes. We anticipated that because of the complexity of pelvic surgery, surgeon-specific volume would be more strongly associated than hospital volume with patient outcomes following ovarian cancer resection. Our analysis did not reveal the large differences in patient outcomes by surgeon volume that have been noted for other cancers *(1,6)*.

We found that the variability in outcomes by surgeon volume was modest and that associations between outcomes and surgeon volume were attenuated after appropriate adjustment for differences in patient mix. We cannot fully explain the specific processes of care that account for the modest volume–outcome associations we observed. However, our finding of a statistically

significant association between higher procedure volume and increased use of postoperative chemotherapy, which persisted after we adjusted for the case mix, suggests that one of the marginally better outcomes of patients treated by high-volume providers as compared with those treated by low-volume providers may be related to their treatment with postoperative chemotherapy. Although our data indicate that chemotherapy delivery may mediate the association between provider procedure volume and survival after ovarian cancer surgery, we cannot determine from this observational study whether the quality of care and postoperative referral processes are better coordinated by high-volume providers than by low-volume providers or, alternatively, whether patients treated by low-volume providers are simply less fit for chemotherapy than those treated by high-volume providers due to unmeasured factors such as performance status or because they were less likely to accept chemotherapy recommendations.

The fact that we did not observe large differences in mortality for patients treated by low- versus high-volume providers suggests that the establishment of minimal procedure volume thresholds that a hospital or surgeon must perform to be accredited as an ovarian cancer surgeon is not easily justified. The Leapfrog group, a consortium of health care purchasers, encourages patients to choose high-volume surgeons *(28)*. Although this strategy makes intuitive sense, our analysis does not indicate that it is supported by the available data for ovarian cancer. In a companion analysis, Earle et al. *(14)* report that the ovarian cancer patients who were treated by gynecologic oncologists and general gynecologists had outcomes that were marginally better than those of patients who were treated by general surgeons. Although it is true that most gynecologic oncologists have high case volumes, some general gynecologists had comparably high case volumes. When considered together, these two studies suggest that high-volume surgeons, many of whom are gynecologic oncologists, achieve marginally better long-term outcomes than low- or intermediate-volume surgeons, many of whom are not gynecologic oncologists.

An alternative strategy to facilitate improvements in the quality of care is to identify providers (either hospitals or physicians) whose patients have especially good or poor outcomes. Close scrutiny of how top-performing surgeons provide care, and detailing the structural systems they have in place that contribute to these superior outcomes may identify strategies that are useful for quality improvement. Giving providers with higher mortality rates feedback about their performance relative to that of their peers and about the strategies used by their more successful colleagues may also lead to better surgical outcomes. Feedback with comparison to peer institutions enables providers to scrutinize those aspects of their practice in need of improvement and is a strategy adopted by organizations such as the National Comprehensive Cancer Network *(21)*. The drawback of such profiling is that it is conditional on volume thresholds. If a hospital or surgeon has only a very few cases, then the small sample size limits the ability to reliably compare outcomes with those of peers.

One issue that is often overlooked in volume–outcome studies is the extent to which especially good or bad outcomes cluster for patients who are cared for at a particular hospital or by a particular surgeon. From a public policy standpoint, it matters enormously whether low-volume providers generally achieve worse outcomes than those of high-volume providers or, alternatively, whether a few high-volume providers whose patients have particularly good outcomes account for the appearance that patients of low-volume providers have generally poor outcomes *(19)*. Our analyses of clustering among hospitals and surgeons in the highest volume categories (Fig. 2) did not reveal substantial variation in outcomes among patients of the high-volume providers.

The limitations of our study are due to the nature of the SEER–Medicare database as a resource for studying the quality of cancer care. First, we could not capture some pertinent information such as the extent to which women had minimal residual disease following ovarian cancer surgery or the volume of their residual disease because this observational dataset lacks the clinical detail that is available from medical record review. Second, our estimates of surgical volume were based on the number of surgeries performed among the Medicare population. Although this method has been validated and Medicare case volume appears to statistically correlate well with total volume *(1)*, it is conceivable that an individual hospital or surgeon who performed a high volume of surgeries in young or HMO-covered patients but only rarely performed procedures among patients in the Medicare population could be misclassified in our analysis *(5)*. Moreover, UPINs submitted on Medicare claims do not perfectly capture surgeon-specific volume because UPINs are missing on some claims. It is because of this potential for misclassification of volume that our study cannot be used to make inferences about the quality of care at a particular institution or in the hands of a particular surgeon. Third, although the SEER cohort is population based, our results may have limited generalizability because our study cohort was restricted to Medicare-eligible patients who were aged 65 years or older and who were not enrolled in an HMO at diagnosis. Fourth, we considered overall mortality rather than disease-specific outcomes such as disease-free survival or cancer-specific mortality because neither recurrence nor cause-specific mortality is reliably captured by SEER–Medicare data. A volume–outcome analysis by Meyerhardt et al. *(22,23)* that used clinical trial data and therefore did permit evaluation of disease specific outcomes found that observed differences reflected variation in patient mix across providers. Fifth, although length of hospital stay is a reliably measured outcome, this metric is not a reliable indicator for quality of care because hospital practices vary with community resources, such as the ability to provide home care. Nevertheless, our finding of an inverse association between the length of hospital stay and procedure volume suggests that high-volume providers are efficient in that they are able to achieve low mortality rates without prolonged hospitalizations. Although we measured ostomy rates because this is an outcome dreaded by patients, there are no data in ovarian cancer relating this outcome to either quality of life or quality of surgical debulking. Finally, the relatively small size of our cohort limited our ability to identify small differences in outcomes that may be associated with provider volume.

Our analysis demonstrates that, throughout the 1990s, the elderly women in our cohort had ovarian cancer surgery performed at many different hospitals, often by surgeons who performed only an occasional resection. We found no evidence that the care for these patients was concentrated among a small group of providers. Although chronologic age and clinical stage dwarfed both hospital and procedure volume as factors associated with outcomes of care, even the modest volume–outcome variations we observed merit further scrutiny to understand the underlying mechanisms that enable higher-volume hospitals and surgeons to achieve more favorable surgical outcomes.

References

(1) Bach PB, Cramer LD, Schrag D, Downey RJ, Gelfand SE, Begg CB. The influence of hospital volume on survival after resection for lung cancer. N Engl J Med 2001;345:181–8.

(2) Begg CB, Cramer LD, Hoskins WJ, Brennan MF. Impact of hospital volume on operative mortality for major cancer surgery. JAMA 1998;280: 1747–51.

(3) Birkmeyer JD, Siewers AE, Finlayson EV, Stukel TA, Lucas FL, Batista I, et al. Hospital volume and surgical mortality in the United States. N Engl J Med 2002;346:1128–37.

(4) Halm EA, Lee C, Chassin MR. Is volume related to outcome in health care? A systematic review and methodologic critique of the literature. Ann Intern Med 2002;137:511–20.

(5) Schrag D, Cramer LD, Bach PB, Cohen AM, Warren JL, Begg CB. Influence of hospital procedure volume on outcomes following surgery for colon cancer. JAMA 2000;284:3028–35.

(6) Birkmeyer JD, Stukel TA, Siewers AE, Goodney PP, Wennberg DE, Lucas FL. Surgeon volume and operative mortality in the United States. N Engl J Med 2003;349:2117–27.

(7) National Cancer Institute. Surveillance, epidemiology, and end results (SEER)-Medicare. Available at: http://healthservices.cancer.gov/seermedicare. [Last accessed: August 1, 2005.]

(8) National Cancer Institute. Surveillance, epidemiology, and end results (SEER). Available at: http://seer.cancer.gov. [Last accessed: August 1, 2005.]

(9) Percy C, Van Holton V, Muir CE. International classification of diseases for oncology. 2nd ed. Geneva (Switzerland): World Health Organization; 1990.

(10) 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.

(11) Sonoda Y. Management of early ovarian cancer. Oncology (Huntingt) 2004;18:343–56; discussion 358, 361–2.

(12) Potosky AL, Riley GF, Lubitz JD, Mentnech RM, Kessler LG. Potential for cancer related health services research using a linked Medicare-tumor registry database. Med Care 1993;31:732–48.

(13) International Classification of Diseases, 9th Revision (ICD-9-CM). Chicago (IL): American Medical Association Publications; 2000.

(14) Earle CC, Schrag D, Neville BA, Yabroff KR, Topor M, Fahey A, et al. Effect of surgeon specialty on processes of care and outcomes for ovarian cancer patients. J Natl Cancer Inst 2006;98:172–80.

(15) Current Procedural Terminology 1999. Chicago (IL): American Medical Association Publications; 1999.

(16) Beebe M, American Medical Association. CPT 2005: current procedural terminology. 4th ed. Chicago (IL): American Medical Association; 2004.

(17) Greene FL, Cancer. AJCo. AJCC Cancer Staging Manual. 6th ed. New York; London (UK): Springer; 2002.

(18) 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; discussion 1081–90.

(19) Panageas KS, Schrag D, Riedel E, Bach PB, Begg CB. The effect of clustering of outcomes on the association of procedure volume and surgical outcomes. Ann Intern Med 2003;139:658–65.

(20) Liang KY, Zeger SL. Longitudinal data-analysis using generalized linear-models. Biometrika 1986;73:13–22.

(21) National Comprehensive Cancer Centers Practice Guidelines. Available at: http://www.nccn.org/professionals/physician_gls/default.as. [Last accessed: August 1, 2005.]

(22) Meyerhardt JA, Catalano PJ, Schrag D, Ayanian JZ, Haller DG, Mayer RJ, et al. Association of hospital procedure volume and outcomes in patients with colon cancer at high risk for recurrence. Ann Intern Med 2003;139:649–57.

(23) Meyerhardt JA, Tepper JE, Niedzwiecki D, Hollis DR, Schrag D, Ayanian JZ, et al. Impact of hospital procedure volume on surgical operation and long-term outcomes in high-risk curatively resected rectal cancer: findings from the Intergroup 0114 Study. J Clin Oncol 2004;22:166–74.

(24) Kumpulainen S, Grenman S, Kyyronen P, Pukkala E, Sankila R. Evidence of benefit from centralised treatment of ovarian cancer: a nationwide population-based survival analysis in Finland. Int J Cancer 2002;102:541–4.

(25) Tingulstad S, Skjeldestad FE, Hagen B. The effect of centralization of primary surgery on survival in ovarian cancer patients. Obstet Gynecol 2003;102:499–505.

(26) Ioka A, Tsukuma H, Ajiki W, Oshima A. Influence of hospital procedure volume on ovarian cancer survival in Japan, a country with low incidence of ovarian cancer. Cancer Sci 2004;95:233–7.

(27) Bristow RE, Zahurak ML, del Carmen MG, Gordon TA, Fox HE, Trimble EL, et al. Ovarian cancer surgery in Maryland: volume-based access to care. Gynecol Oncol 2004;93:353–60.

(28) The Leapfrog Group. Available at: http://www.leapfroggroup.org/.

Notes

We thank the Applied Research Program, NCI; the Office of Information Services, and the Office of Strategic Planning, CMS; Information Management Services Inc.; and the Surveillance, Epidemiology, and End Results (SEER) Program tumor registries in the creation of the SEER–Medicare database used in this study. Drs. Diane Bodurka and Michael Carney provided valuable input into the definition of surgical procedures used in this analysis.

Supported by the Health Services and Economics Branch, Division of Cancer Control and Population Sciences and the Cancer Therapy and Evaluation Branch of the National Cancer Institute. The study authors were solely responsible for the data collection, analysis, and interpretation, manuscript preparation, and the decision to submit the manuscript for publication.

Manuscript received April 15, 2005; revised November 18, 2005; accepted December 6, 2005.