Original Article

Variations in Surgical Spending Within Hospital Systems for Complex Cancer Surgery

Adrian Diaz, MD, MPH [1,2,3]; Karan R. Chhabra, MD, MSc[2,3,4]; Justin B. Dimick, MD, MPH[3,5]; and Hari Nathan, MD, PhD[3,5]

BACKGROUND: Approximately 70% of hospitals today are part of larger health systems. Proponents of hospital consolidation tout its potential to reduce health spending and improve outcomes, but to the authors' knowledge the available evidence has suggested that this promise is unrealized. Variations in costs and outcomes within systems may highlight opportunities for collaborative quality improvement and practice standardization. To assess this potential, the authors sought to measure variations in episode spending within and across hospital systems among Medicare beneficiaries undergoing complex cancer surgery. **METHODS:** Using 100% Medicare claims data, the authors identified fee-for-service Medicare patients who were undergoing elective pancreatectomy, lung resection, or colectomy for cancer from 2014 through 2016. Risk-adjusted, price-standardized payments for the surgical episode from admission through 30 days after discharge were calculated. The authors then assessed the reliability-adjusted variations at the hospital and system levels. **RESULTS:** Average episode payments varied nearly as much within hospital systems for pancreatectomy ($1946 between the lowest and highest spending systems; 95% CI, $1910-$1972), lung resection ($625 between the lowest and highest spending systems; 95% CI, $621-$630), and colectomy ($813 between the lowest and highest spending systems; 95% CI, $809-$817) as they did between the lowest and highest spending hospitals (pancreatectomy: $2034; lung resection: $1789; and colectomy: $770). For pancreatectomy, this variation was driven by index hospitalization spending whereas both index hospitalization and postacute care use drove variations for lung resection and colectomy. **CONCLUSIONS:** In this analysis of Medicare patients undergoing complex cancer surgery, wide variations in surgical episode spending were noted both within and across hospital systems. System leaders may seek to better understand variations in practices among their hospitals to standardize care and reduce variations in outcomes, use, and costs. ***Cancer*** **2021;127:586-597**.

KEYWORDS: cancer surgery, colectomy, health systems, pancreatectomy, pneumonectomy, quality, spending.

INTRODUCTION

The accelerating pace of hospital consolidation over the last 20 years has resulted in approximately 70% of hospitals today belonging to larger health systems.[1,2] Proponents of hospital mergers tout their potential to improve outcomes and reduce health spending by standardizing care, eliminating redundancy, and achieving economies of scale.[3] However, to our knowledge, there is conflicting evidence regarding whether expanding hospital systems are reducing health spending.[4-7] Furthermore, wide variations in surgical episode expenditures within health systems have been observed.[8,9] These observations call into question whether systems are leveraging their collective volume and experience to standardize care and maximize efficiencies If care patterns within a health system vary significantly, this suggests that the system may have failed to address gaps in practice standardization and clinical quality across its affiliated hospitals.[9]

Because cancer care can be high risk and often multidisciplinary, it may lend itself well to being optimized through hospital consolidation. Many systems have taken the first step of using their expanding referral networks to concentrate volumes and improve quality within these service lines. However, inconsistencies in quality may persist. Recent research has demonstrated that for hospitals performing hip replacements, discharge patterns vary nearly as much within systems as between systems, indicating that inconsistencies in care coordination persist within systems.[8] Variations in surgical expenditures, which can be reliably measured, are a useful marker for understanding the degree of standardization of surgical quality and care pathways. For example, systems may control episode spending by internally referring patients to their lowest spending hospitals or disseminating low-cost centers' care patterns throughout

Corresponding Author: Adrian Diaz, MD, MPH, Institute for Healthcare Policy and Innovation Clinician Scholars Program, University of Michigan, 2800 Plymouth Rd, Bldg 14, Rm G100, Ann Arbor, MI 48109 (Adriandi@med.umich.edu).

[1]Department of Surgery, The Ohio State University, Columbus, Ohio; [2]Institute for Healthcare Policy and Innovation Clinician Scholars Program, University of Michigan, Ann Arbor, Michigan; [3]Center for Healthcare Outcomes and Policy, University of Michigan, Ann Arbor, Michigan; [4]Department of Surgery, Brigham and Women's Hospital, Boston, Massachusetts; [5]Department of Surgery, University of Michigan, Ann Arbor, Michigan

This analysis has been accepted for presentation at the 2020 Society for Surgical Oncology Annual Meeting (postponed because of the coronavirus disease 2019 pandemic).

Additional supporting information may be found in the online version of this article.

DOI: 10.1002/cncr.33299, **Received:** June 4, 2020; **Revised:** September 7, 2020; **Accepted:** October 10, 2020, **Published online** November 3, 2020 in Wiley Online Library (wileyonlinelibrary.com)

their affiliates.[10] However, to our knowledge, whether systems actually are leveraging their position to reduce variations in surgical episode spending for high-risk cancer surgery remains unknown.

In the current study, we explored surgical spending variations within and across hospital systems in the United States using Medicare data regarding patients undergoing surgical resections of the pancreas, lung, or colon for cancer. We chose these procedures because they range in both frequency and complexity. We hypothesized significant variations in surgical episode spending on high-risk cancer procedures both within and across health systems.

MATERIALS AND METHODS

Data Source and Population

We used 100% claims data from the Medicare Provider Analysis and Review (MEDPAR) file for cancer resections performed from calendar years 2014 through 2016 at nonfederal acute care hospitals. We used procedure codes for colon resection, lung resection, and pancreatic resection from the *International Classification of Diseases, Ninth* Revision (ICD-9) and *International Classification of Diseases, 10th Revision Procedure Coding System* (ICD-10-PCS) from the MEDPAR file, with confirmatory Current Procedural Terminology codes from the Medicare Carrier File, to define the cohort. We included fee-for-service Medicare patients aged 66 to 99 years with continuous coverage for 3 months before and 6 months after the surgical procedure of interest. We excluded nonelective admissions, hospitals with <10 fee-for-service Medicare cases across all 3 years, hospitals that were not participating in systems, and patients with Medicare Advantage coverage. In addition, we excluded any systems with <2 hospitals represented in our cohort. Hospital identifiers from the MEDPAR file were linked to the American Hospital Association Annual Survey of Hospitals for each corresponding year, which provided system identifiers and hospital characteristics.

Outcomes

The primary outcomes in the current study were Medicare's 30-day episode payments ("spending"), with specific attention to the total episode payment, index hospitalization, and postacute care component of the total payment. These were derived from the MEDPAR, carrier, outpatient, and home health agency files. "Episodes" encompassed the index procedure with associated hospitalization and postacute care services, physician services, readmissions, and outpatient care up to 30 days after discharge. We defined "institutional postacute care" as discharge to a facility for skilled nursing, inpatient rehabilitation, intermediate care, or long-term care.[11] We used price standardization methods that previously were described to adjust for intended differences in Medicare payment rates (by year, wage index, and graduate medical education expenses).[12-15] In addition, we have provided more detailed methodology regarding our price standardization in the Supporting Materials.

Definitions

Because centralizing surgical volume at the highest value centers is one potential strategy for minimizing variations and improving outcomes, we measured the procedural volume at each system's hub relative to the system as a whole. We identified "hubs" as the hospitals with the highest Medicare volume for each procedure within each system. We calculated the volume "concentration" by dividing the hub's case volume by the total system case volume for each procedure for each system.[16]

Statistical Analysis

Patient-level episode payments were winsorized to the 1st and 99th percentiles to limit the influence of extreme outliers. Winsorization recodes extreme outliers to less extreme values, thereby improving model fit while preserving the underlying signal and without deleting observations.[17] Our use of winsorization did not remove extreme values; rather, it reset them to less extreme values. Thus, very extreme outliers (eg, payments at the 99.9th percentile) would be reset to the 99th percentile, and hospitals with many extreme outliers would continue to have higher average payments despite winsorization. Outliers were equally present at the high and low ends of the distribution. Winsorization occurred after price standardization. Overall, approximately 2.0% of the data were winsorized (pancreas: 1.9%; lung: 2.0%; and colon: 2.0%). We then fit multilevel linear regression models with hospital-level and system-level random effects to estimate the average total episode payments. Risk adjustment controlled for patient age (as a quadratic term), sex, and Elixhauser comorbidities. In addition to this, we adjusted for ICD-9 and ICD-10-PCS principal diagnosis and procedure codes (as indicator variables) to capture differences in patient condition and surgical approach.

We then used the models to calculate risk-adjusted and reliability-adjusted average payments at the hospital and system levels. Reliability adjustment was used to reduce statistical "noise" and create more accurate hospital rankings.[18-20] This technique filters out statistical noise by shrinking the observed rate toward the average rate. This problem can be conceptualized by imagining a hospital with a mortality rate of 0, but only 5 cases. It is highly likely that, if the hospital were to perform 100 more cases, the mortality rate would no longer be zero. Reliability uses hierarchical modeling techniques adjusted estimates based on sample size variation so that deviations from average are much more likely to represent true deviations rather than statistical noise.

For the current study, we used a 2-level model with the hospital as the first level and the system as the second level. Using postestimation commands, we created empirical Bayes estimates of the random effect of each hospital and system. These random effects represented the risk-adjusted and reliability-adjusted "signal." Models that treated random effects as independent, exchangeable, and unstructured were compared and no differences were noted with regard to the outcomes. In the final model, random effects were treated independently because it was unlikely that hospitals within a system were coordinating with one another in a meaningful way based on our hypothesis.

We calculated hospital average spending by adding the national average total episode payment to the hospital's and relevant system's best linear unbiased predictors. We calculated system average spending by adding the national average total episode payment to the best linear unbiased predictor of the system's random effect. These risk-adjusted and reliability-adjusted estimates were used to divide systems into quintiles of average spending, weighted at the episode level. Payment variations within systems were calculated by comparing the highest spending and lowest spending hospital's risk-adjusted and reliability-adjusted averages.

Postoperative outcomes and discharge destination risk-adjusted rates were derived using marginal means in logistic regression models in which the outcome was treated as a categorical variable. All models were adjusted for patient age, sex, and race; 27 Elixhauser comorbidities; overall time trends; and hospital characteristics. In addition to this, we adjusted for ICD-9 and ICD-10-PCS principal diagnosis and procedure codes (as indicator variables) to capture differences in patient condition and surgical approach. All outcomes and discharge destination were reported as risk-adjusted rates.

All analyses were performed using Stata statistical software (version 16; StataCorp LLC, College Station, Texas).

The current study was deemed exempt by the University of Michigan institutional review board.

RESULTS

Hospital and System Characteristics

We identified 57,458 Medicare patients who underwent either pancreatic, lung, or colon resection from 2014 to 2016 (Table 1). For the patients who underwent pancreatectomy, 5415 episodes occurred at 322 hospitals within 95 systems. Average total episode payments ranged from $31,481 (first quintile) to $33,427 (fifth quintile) across systems. For patients who underwent lung resection, a total of 23,285 episodes occurred at 563 hospitals within 111 systems. Average total episode payments ranged from $20,700 (first quintile) to $21,325 (fifth quintile) across systems. For patients who underwent colectomy, 28,849 episodes occurred at 900 hospitals within 160 systems. Average total episode payments ranged from $18,417 (first quintile) to $19,513 (fifth quintile) across systems. For each operation, patients' age, sex, race, and Elixhauser comorbidity counts were similar across all quintiles. The mean number of hospitals per system ranged from 10 for pancreatectomy to 17 for colectomy. For each surgery, the percentage of patients treated at a for-profit hospital was highest in the highest spending quintile.

Although procedures were analyzed independently, there were hospitals and hospital systems that were low cost for >1 procedure. Specifically, 40 hospitals had lower than mean spending for all 3 procedures studied and 90 hospitals had lower than mean spending for 2 of the 3 procedures. Among hospital systems, 49 had lower than mean spending for all 3 procedures studied and 99 had lower than mean spending for 2 of the 3 procedures.

Variations Across Systems

The difference in total episode payments for pancreatectomy between the lowest and highest spending quintiles of systems was $1946 (95% CI, $1910-$1972), or 6.2% of total episode spending in the lowest quintile (Fig. 1). The largest component of this variation (39%) was accounted for by index hospitalization spending, followed by postacute care spending (27%) (Table 2). The lowest spending quintile had an average risk-adjusted serious complication rate of 8.6% compared with 15.9% in the

TABLE 1. Health System Characteristics by Quintile of Spending for Each Surgery

		Quintile of Average System Spending					
	Overall	1	2	3	4	5	*P*
	Pancreatectomy						
Average total episode payment	$32,264	$31,481	$31,901	$32,114	$32,426	$33,427	
No. of cases	5415	1096	1109	1088	1042	1080	
Mean age (SD), y	74.3 (5.9)	74.8 (6.0)	74.2 (5.7)	74.2 (6.0)	74.3 (6.0)	74.1 (5.7)	.068
Male	51%	48%	53%	51%	51%	52%	.213
White	89%	90%	91%	87%	89%	90%	.006
Black	6%	7%	5%	8%	4%	7%	.005
No. of comorbidities (%)							
0	232 (4)	48 (4)	51 (5)	42 (4)	43 (4)	46 (4)	.935
1	730 (13)	114 (10)	159 (14)	133 (12)	151 (15)	161 (15)	.008
>2	4609 (83)	934 (85)	899 (81)	913 (84)	848 (81)	873 (81)	.020
No. of systems	95	13	13	34	19	16	
Mean no. of cases per system	140	163	156	71	136	108	<.001
Mean no. of hospitals per system	10	10	5	3	7	20	<.001
Average annual volume at hub	97	173	119	67	63	55	<.001
Average no. of beds at hub	894	945	1247	752	718	768	<.001
Case concentration	76%	82%	76%	87%	65%	63%	<.001
No. of hospitals	322	61	42	61	77	81	
Mean no. of cases per hospital	76	125	96	56	53	36	<.001
Patients treated at teaching hospital	65%	67%	67%	81%	66%	51%	<.001
Patients treated at urban hospital	97%	100%	97%	100%	95%	94%	<.001
Patients treated at for-profit hospital	9%	0%	2%	0%	10%	32%	<.001
	Lung Resection						
Average total episode payment	$20,973	$20,700	$20,839	$20,929	$21,079	$21,325	
No. of cases	23,285	4880	4447	4688	4665	4605	
Mean age (SD), y	74.0 (5.6)	73.8 (5.5)	73.9 (5.6)	74.1 (5.6)	74.1 (5.6)	74.0 (5.5)	.031
Male	47%	48%	48%	45%	46%	48%	.004
White	91%	91%	90%	91%	91%	93%	<.001
Black	6%	6%	8%	5%	5%	4%	<.001
No. of comorbidities (%)							
0	1377 (6)	277 (6)	219 (5)	286 (6)	247 (7)	243 (5)	<.001
1	4005 (17)	803 (17)	697 (16)	854 (18)	918 (20)	717 (16)	<.001
>2	17,971 (77)	3800 (78)	3531 (79)	5548 (76)	3400 (73)	3645 (79)	<.001
No. of systems	111	19	19	28	34	11	
Mean no. of cases per system	489	511	412	352	228	946	
Mean no. of hospitals per system	13	10	11	5	4	29	<.001
Average annual volume at hub	166	150	133	192	197	92	<.001
Average bed size at hub	792	894	614	690	938	771	<.001
Case concentration	59%	47%	53%	65%	70%	42%	<.001
No. of hospitals	563	99	108	97	110	149	
Mean no. of cases per hospital	103	91	79	155	139	49	
Patients treated at teaching hospital	40%	46%	28%	52%	52%	19%	<.001
Patients treated at urban hospital	97%	97%	95%	99%	96%	95%	<.001
Patients treated at for-profit hospital	13%	4%	1%	2%	4%	56%	<.001

TABLE 1. *Continued*

		Quintile of Average System Spending					
	Overall	1	2	3	4	5	*P*
	Colectomy						
Average total episode payment	$18,919	$18,417	$18,638	$18,890	$19,161	$19,513	
No. of cases	28,849	5987	5960	6409	5546	5910	
Mean age (SD), y	76.8 (7.3)	77.0 (7.3)	76.6 (7.3)	76.7 (7.3)	76.9 (7.4)	76.8 (7.3)	.009
Male	46%	46%	47%	46%	47%	47%	.412
White	89%	88%	89%	89%	87%	89%	<.001
Black	7%	8%	7%	7%	9%	7%	.002
No. of comorbidities (%)							
0	2071 (7)	396 (7)	446 (8)	431 (7)	347 (6)	448 (8)	.018
1	5067 (17)	1038 (17)	1020 (17)	1047 (16)	951 (17)	1005 (17)	.624
>2	22,711 (76)	4533 (76)	4494 (75)	4931 (77)	4248 (77)	4457 (75)	.174
No. of systems	160	25	36	37	40	22	
Mean no. of cases per system	548	432	385	462	231	1217	<.001
Mean no. of hospitals per system	17	12	10	12	8	42	<.001
Average annual volume at hub	78	84	77	64	83	84	<.001
Average bed size at hub	663	624	711	610	614	800	<.001
Case concentration	47%	42%	46%	46%	53%	46%	<.001
No. of hospitals	900	201	177	210	187	215	
Mean no. of cases per hospital	48	48	53	44	55	43	<.001
Patients treated at teaching hospital	23%	21%	27%	23%	30%	16%	<.001
Patients treated at urban hospital	94%	94%	92%	95%	95%	94%	<.001
Patients treated at for-profit hospital	15%	3%	1%	5%	11%	57%	<.001

Quintiles of system spending were generated from risk-adjusted and reliability-adjusted estimates of system average payments and weighted at the episode level. Payments were price-standardized and winsorized to the 1st and 99th percentiles. Hospital and system characteristics were weighted at the episode level to reflect differences in hospital and system volumes. Case concentration was calculated as the hub's case volume divided by the total system case volume for each system. Significance testing also was performed at the episode level using univariate statistics (analysis of variance) as appropriate.

highest spending quintile ($P < .001$). Fewer patients were discharged home in the highest spending quintile versus the lowest spending quintile (34% vs 45%) (Table 3).

The difference in total episode payments for lung resection between the lowest and highest spending quintiles of systems was $625 (95% CI, $621-$630), or 3.0% of total episode spending in the lowest spending quintile (Fig. 2). For colectomy, this difference was $813 (95% CI, $809-$817), or 4.4% of total episode spending in the lowest spending quintile. It is interesting to note that for 39 systems (24.4%), every hospital had higher average episode spending than the national average (Fig. 3). For both surgical procedures, postacute care spending explained the largest component of the variation in spending (37% for lung and 47% for colon), followed by index hospitalization payments. For lung resection, the highest spending systems discharged approximately 54% of patients home, compared with 63% in the lowest spending systems. Similarly, for colectomy, the highest-spending systems discharged approximately 66% of patients home, compared with 61% in the lowest spending systems. There was no difference noted with regard to postoperative complications, mortality, or readmissions related to episode spending between the lowest spending and highest spending quintiles for lung resection. However, for colectomy, postoperative complications were found to be 1.8% higher in the highest spending versus the lowest spending quintile. This difference in quality may have contributed to the 15% variation in index hospitalization spending and the 25% variation in physician payment variation (Table 3).

Average total episode spending at the hubs of the lowest spending systems was lower than spending at the hubs of the highest spending systems (Table 4). For pancreatectomy, spending was $5349 less at the hubs of the lowest spending systems compared with the hubs of the highest spending systems. Furthermore, case volume (174 cases vs 64 cases), concentration (85% vs 64%), and number of hospital beds (1027 beds vs 591 beds) all were greater at the hubs of the lowest spending

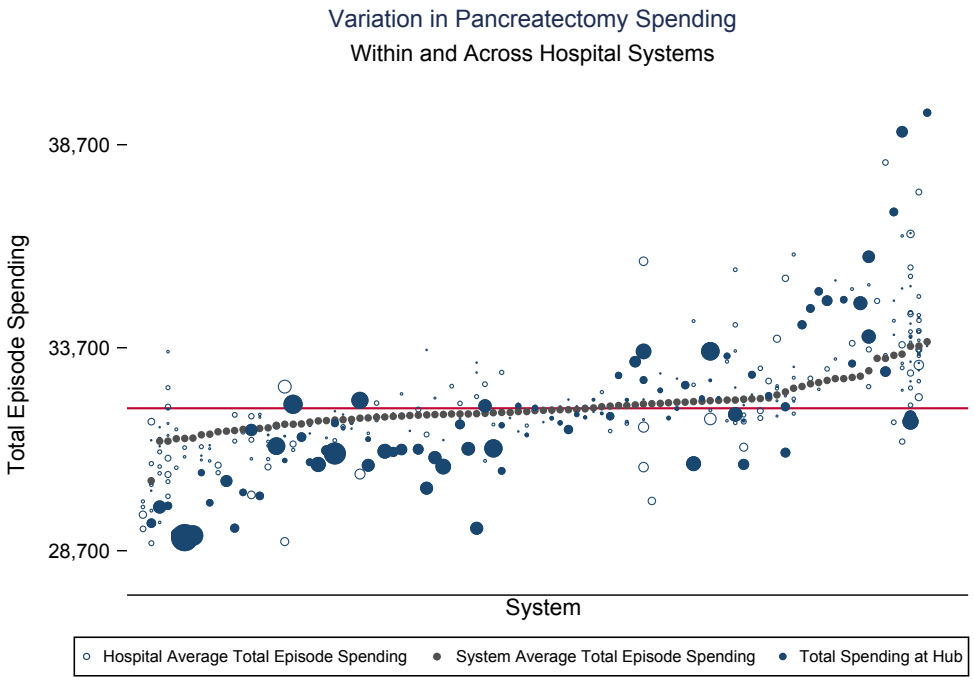

Figure 1. Variations in pancreatectomy spending within and across health systems. System and hospital average payments were price-standardized, winsorized to the 1st and 99th percentiles, and risk-adjusted and reliability-adjusted. The horizontal line represents the national average 30-day episode payment for pancreatectomy ($32,264). The small black dots represent the average episode payments at the system level; systems are arrayed from the lowest to the highest average spending. The bubbles represent average episode payments at the hospital level; larger bubbles represent higher volume hospitals. Bubbles within a vertical column represent hospitals within the same system. The solid circles represent the system hub (the hospital within each system with the highest volume).

systems compared with the hubs of the highest spending systems (all $P < .001$). Finally, hubs from the lowest cost systems discharged patients home at a higher rate than hubs from the highest spending systems (48% vs 28%; $P < .001$).

For lung resection and colectomy, spending was $2833 and $1498, respectively, less at the hubs of the lowest spending systems compared with the hubs of the highest spending systems. Contrary patterns observed for pancreatectomy, case volume, concentration, and hospital beds were either lower or not significantly different at the hubs of the lowest spending systems compared with the hubs of the highest spending systems (Table 4). Hubs from the lowest spending systems discharged patients home at a higher rate than hubs from the highest spending systems (69% vs 41% for lung resection and 69% vs 56% for colectomy [both $P < .001$).

Variations Within Systems

Risk-adjusted and reliability-adjusted episode payment variations for all 3 surgeries were greater within systems than across systems (Table 2). The variation between the lowest spending and highest spending hospitals within a system was, on average, $2034 (interquintile range, $772-$5682) for pancreatectomy, $1789 (interquintile range, $465-$4305) for lung resection, and $770 (interquintile range, $211-$1575) for colectomy. The index hospitalization spending largely explained the within-system variations in hospital spending for all 3 surgeries. It is interesting to note that greater variation between the lowest spending and highest spending hospitals within the system was associated with a modestly higher system average total episode payment.

Case Concentration

Average case concentration at the system hub ranged from 47% for colectomy to 76% for pancreatectomy. Greater concentration was associated with lower spending for pancreatectomy. Pancreatectomy case concentration ranged from 82% in the lowest spending quintile to 63% in the highest spending quintile. There was a difference of $934 (95% CI, $932-$935) between the least centralized and most centralized systems (see Table S1). Conversely, for colectomy, the case concentration varied

TABLE 2. Variations in Spending Across and Within Systems for Each Surgery

	Variations Across Systems				Variations Within System			
		System Quintile				System Quintile		
	Overall	1	5	Total Episode Payment Difference	Overall	1	5	Total Episode Payment Difference
Pancreatectomy								
Episode payments								
Average total episode	$32,264	$31,481	$33,427	100%	$2034	$722	$5682	100%
Average index hospitalization	$20,211	$19,720	$20,479	39%	$1302	$462	$3636	64%
Average physician	$5760	$5620	$5990	19%	$244	$87	$682	12%
Average PAC	$3942	$3846	$4371	27%	$386	$137	$1080	19%
Average readmission	$2352	$2295	$2607	16%	$102	$36	$284	5%
Lung resection								
Episode payments								
Average total episode	$20,973	$20,700	$21,326	100%	$1789	$465	$4305	100%
Average index hospitalization	$13,902	$13,721	$13,947	36%	$966	$251	$2325	54%
Average physician	$3592	$3545	$3658	18%	$286	$74	$689	16%
Average PAC	$2418	$2387	$2618	37%	$447	$116	$1076	24%
Average readmission	$1061	$1047	$1053	10%	$89	$23	$215	5%
Colectomy								
Episode payments								
Average total episode	$18,918	$18,559	$19,372	100%	$770	$211	$1575	100%
Average index hospitalization	$12,450	$12,676	$12,798	15%	$408	$112	$835	53%
Average physician	$3120	$2987	$3190	25%	$123	$34	$252	16%
Average PAC	$2359	$1961	$2343	47%	$185	$51	$378	24%
Average readmission	$990	$936	$1042	13%	$54	$15	$110	7%

Abbreviation: PAC, postacute care facility (eg, skilled nursing facility, rehabilitation facility, etc).
Quintiles of system spending were generated from risk-adjusted and reliability-adjusted estimates of system average payments and weighted at the episode level. Within-system variations were generated by subtracting the lowest spending hospital payments from the highest spending hospital payments for each system. Payments were price-standardized and winsorized to the 1st and 99th percentiles.

from 42% to 53% but with no systematic variations noted across quintiles of spending (Table 1). There was a difference of only $80 (95% CI, $79-$81) observed between the least centralized and most centralized systems. Similar findings were observed for lung resection: there was a difference of only $104 (95% CI, $104-$110) noted between the least centralized and most centralized systems.

DISCUSSION

Proponents of hospital consolidation have highlighted its potential to reduce health spending by standardizing care, eliminating redundancy, and achieving economies of scale. To our knowledge, whether these advantages have been realized for patients, payers, and the systems themselves is unknown. Variation in surgical expenditures is one measurable indicator of the degree of quality and care standardization achieved by a system. We would expect that well-functioning systems should both reduce average expenditures and reduce the variations in expenditures within their systems. Furthermore, because complications, quality of care, and care coordination often are correlated with costs of care,[21-23] variations in spending within and across systems may signal opportunities for quality improvement. In the current analysis, among Medicare beneficiaries undergoing oncologic resections of the pancreas, lung, or colon, we observed significant variations in surgical episode spending both within and across hospital systems. Compared with the across-system variation, the degree of within-system variation may reflect the degree to which hospital systems have standardized care in particular service lines.

For the cancer resection procedures studied, there were modest variations in spending noted between systems, ranging from 3.0% for lung resection to 6.2% for pancreatectomy. However, given the number of procedures performed, the aggregate expenditures associated

TABLE 3. Postoperative Outcomes and Discharge Destinations for the Highest and Lowest Spending Systems and Hospitals for Each Surgery

		Variations Across Systems			Variations Within System		
		System Quintile			System Hospital		
	Overall	1	5	*P*	Lowest Spending	Highest Spending	*P*
Pancreatectomy							
Postoperative outcomes							
Serious complications	12.4	8.6	15.9	<.001	11.1	12.7	.277
30-d mortality	4.8	5.5	5.5	.992	5.0	4.0	.353
Readmission	24.0	22.0	29.1	.005	23.9	23.4	.756
Discharge destination							
Average across system							
Home	39%	45%	34%		46%	32%	
Institutional PAC	22%	22%	28%		17%	26%	
Home health	35%	28%	31%		35%	37%	
Other	5%	5%	7%		3%	5%	
Lung resection							
Postoperative outcomes							
Serious complications	8.0	6.9	9.4	<.001	6.8	8.1	.003
30-d mortality	2.3	2.2	2.8	.082	2.0	1.8	.493
Readmission	11.5	10.7	11.8	.156	11.2	11.2	.975
Discharge destination							
Average across system							
Home	57%	63%	54%		61%	60%	
Institutional PAC	12%	9%	16%		10%	12%	
Home health	29%	25%	27%		28%	27%	
Other	2%	2%	3%		2%	2%	
Colectomy							
Postoperative outcomes							
Serious complications	4.9	3.9	5.6	<.001	4.7	5.4	.078
30-d mortality	2.5	2.6	2.6	.901	2.2	2.1	.839
Readmission	12.3	10.7	13.6	.000	11.5	13.2	.025
Discharge destination							
Average across system							
Home	63%	66%	61%		65%	61%	
Institutional PAC	16%	13%	17%		15%	16%	
Home health	18%	19%	18%		18%	20%	
Other	3%	3%	4%		2%	3%	

Abbreviation: PAC, postacute care facility (eg, skilled nursing facility, rehabilitation facility, etc).
For postoperative outcomes, the risk-adjusted rates were derived using marginal means in logistic regression models in which outcome was treated as a categorical variable. All models were adjusted for patient age and sex and 27 Elixhauser comorbidities, overall time trends, and hospital characteristics.

with such variations are significant. For comparison, the Centers for Medicare and Medicaid Services's bundled payment programs, including the Acute Care Episode Demonstration, the voluntary Bundled Payments for Care Improvement initiative, and the mandatory Comprehensive Care for Joint Replacement model, have yielded mixed results, ranging from no change in spending to 3.9% savings for select episodes such as joint replacement.[24,25] With that in mind, the variations in surgical episode spending observed within and across hospitals systems present sizable opportunities for system leaders to realize significant savings, comparable to those achieved by large-scale alternative payment models.

For pancreatectomy, variations both within and across systems were driven by index hospitalization payments. In addition, the lowest spending systems were found to have significantly lower postoperative complication rates compared with the highest spending systems. Increased centralization of pancreatectomy was associated with lower spending and better quality, as we have reported previously.[16] These findings suggest one strategy that systems may use to optimize outcomes and the costs of complex, less common surgeries such as pancreatectomy: centralize volume and focus quality improvement efforts at 1 or a few referral centers. This strategy is consistent with the robust literature documenting volume-outcome relationships for complex surgical procedures,[26-29] but the implementation of volume-based referral within a hospital system avoids the disincentive of lost revenue that otherwise may prevent low-volume hospitals from turning away such cases.[30]

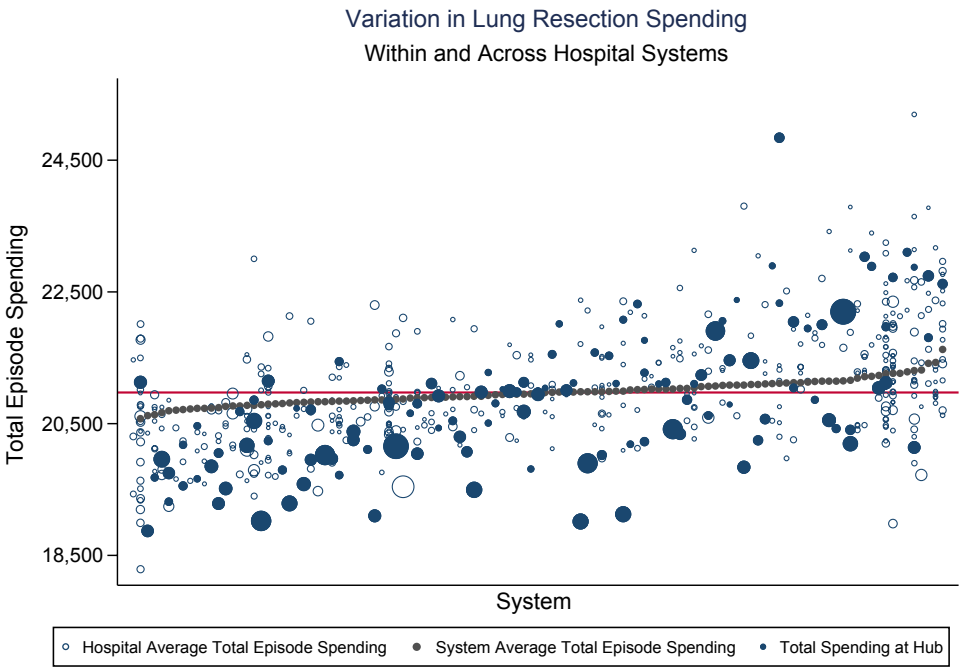

Figure 2. Variations in lung resection spending within and across health systems. System and hospital average payments were price-standardized, winsorized to the 1st and 99th percentiles, and risk-adjusted and reliability-adjusted. The horizontal line represents the national average 30-day episode payment for lung resection ($20,973). The small black dots represent the average episode payments at the system level; systems are arrayed from the lowest to the highest average spending. The bubbles represent average episode payments at the hospital level; larger bubbles represent higher volume hospitals. Bubbles within a vertical column represent hospitals within the same system. The solid circles represent the system hub (the hospital within each system with the highest volume).

Similarly, for more common procedures such as lung and colon resection, variations within systems were driven by index hospitalization payments, suggesting that hospitals within the same system have significant differences in use patterns even around the same surgery, most likely relating to differences in practice patterns, complication rates, or both. However, for both procedures, variations across systems were dominated by differences in postacute care payments. Although complication rates and index hospitalization payments did vary significantly across systems, the relatively larger variations in postacute care payments suggested that optimizing the discharge destination may be an opportunity to control spending. For example, our group has shown that for hospitals performing hip replacements, discharge patterns varied nearly as much within systems as they did across systems, contributing to approximately 86% of spending variations between the lowest and highest spending systems.[10] This variation in discharge destination after joint replacement commonly is believed to be due to idiosyncratic provider preferences rather than unmeasured differences among patients. Although to our knowledge it has not been well studied, variation in postacute care use after cancer resection may be similarly discretionary. In these instances, in which surgeries commonly are performed at various affiliated hospitals within a system, surgeons and hospital leaders have an opportunity to standardize best practices across the health system. In the current study, we observed similar discharge patterns among patients undergoing lung resection and colectomy. As such, interventions aimed as standardizing care, such as the Enhanced Recovery After Surgery (ERAS) protocol and reductions in postacute care use, have not only improved quality but reduced cost.[31-33] Furthermore, for colectomy, nearly 25% of systems had all of their hospitals demonstrate average episode payments that were greater than the national average. This may signal system-wide opportunities for surgeons and health system leaders to ensure uniform standards of care and reduce spending across the entire system. It is important to note that for some of the highest spending systems, nearly all of the hospitals were above the average system spending whereas other high spending systems were found to have some hospitals for which spending was less than average (Figs. 1-3). This observation may signal that for some systems, best practices from low spending hospitals can be disseminated across the

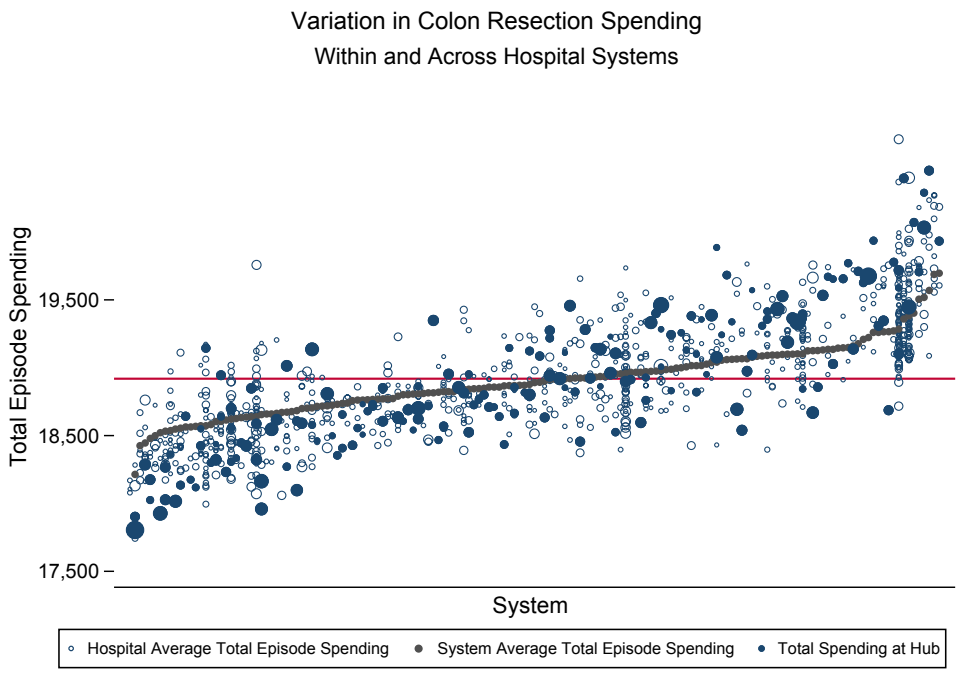

Figure 3. Variations in colectomy spending within and across health systems. System and hospital average payments were price-standardized, winsorized to the 1st and 99th percentiles, and risk-adjusted and reliability-adjusted. The horizontal line represents the national average 30-day episode payment for colectomy ($18,918). The small black dots represent average episode payments at the system level; systems are arrayed from the lowest to the highest average spending. The bubbles represent average episode payments at the hospital level; larger bubbles represent higher volume hospitals. Bubbles within a vertical column represent hospitals within the same system. The solid circles represent the system hub (the hospital within each system with the highest volume).

network, whereas for other systems, best practices may need to be re-evaluated across the network.

The approach used in the current study had several limitations. First, the current analysis did not distinguish systems based on size or geographic spread, nor were we able to adjust for the degree of clinical or financial integration within these systems. Although some may interpret the system as a loose affiliation of hospitals, others may be more deliberate in coordinating providers and services. Second, similar to other studies using administrative data, the findings of the current study are subject to residual confounding due to unmeasured factors such as patient severity, noncoded comorbidities, and sociodemographic factors. Third, the study was not longitudinal in nature in that we were unable to assess whether hospital spending changes after joining a system. Finally, it is important to note that not all institutional postacute care is wasteful, and that reducing the use of these facilities past a certain point may be harmful for patients. However, there is reasonable evidence that postacute care facilities do not improve long-term outcomes for patients with cancer.[34]

Although hospital mergers proceed on the premise that they will achieve better care at lower cost, the data from the current study have demonstrated that wide variations in spending across and within hospital systems for cancer surgery persist. Despite being affiliated in name, systems still may lack the financial alignment to strategically redistribute care. The wide variations in episode spending observed in the current study may indicate that the promise of better care at a lower cost through care coordination has not been fully achieved. Because both private and public payers aim to drive down surgical episode spending, systems will need to better control episode spending to remain competitive. The results of the current study suggested that strategies to curb variations in spending may have to be tailored to different surgeries within each system. For example, some high-risk cancer surgeries may be best suited for centralization, whereas others may be better targets for standardization across the system. As hospital systems continue to evolve in response to broader financial pressures, surgeons and system leaders may be presented with unique opportunities for improving both the quality and cost of cancer care.

TABLE 4. Hub Characteristics from the Lowest and Highest Spending Systems

		Across System Hubs		
		System Quintile		
	Overall	1	5	*P*
Pancreatectomy				
Average total episode	$31,639	$29,274	$34,623	<.001
Hospital volume	96	174	64	<.001
Concentration	76%	85%	64%	<.001
Teaching hospital	82%	93%	81%	<.001
Hospital beds	908	1027	591	<.001
Discharge destination				
Home	37%	48%	28%	<.001
Institutional PAC	21%	20%	26%	
Home health	36%	29%	43%	
Other	4%	4%	3%	
Lung resection				
Average total episode	$20,593	$19,357	$22,190	<.001
Hospital volume	166	154	207	<.001
Concentration	59%	60%	63%	<.001
Teaching hospital	63%	72%	67%	.002
Hospital beds	797	808	1090	<.001
Discharge destination				
Home	60%	69%	41%	<.001
Institutional PAC	10%	7%	16%	
Home health	28%	22%	41%	
Other	2%	2%	2%	
Colectomy				
Average total episode	$18,898	$18,164	$19,662	<.001
Hospital volume	78	89	91	.17
Concentration	47%	47%	44%	<.001
Teaching hospital	46%	41%	40%	.78
Hospital beds	664	676	873	<.001
Discharge destination				
Home	63%	69%	56%	<.001
Institutional PAC	15%	12%	19%	
Home health	19%	16%	23%	
Other	2%	3%	3%	

Abbreviation: PAC, postacute care facility (eg, skilled nursing facility, rehabilitation facility, etc).

Quintiles of system spending were generated from risk-adjusted and reliability-adjusted estimates of hospital average payments at the hub for each system. Payments were price-standardized and winsorized to the 1st and 99th percentiles. Case concentration was calculated as the hub's case volume divided by the total system case volume for each system. Significance testing also was performed at the episode level using univariate statistics as appropriate.

FUNDING SUPPORT

Justin B. Dimick received grant funding from the National Institutes of Health (R01AG039434). Hari Nathan has received funding from the Agency for Healthcare Research and Quality (K08HS024763) and the National Institutes of Health (R01AG039434).

CONFLICT OF INTEREST DISCLOSURES

Adrian Diaz has received funding from the University of Michigan Institute for Healthcare Policy and Innovation Clinician Scholars Program. Karan R. Chhabra has received payments from Blue Cross Blue Shield of Massachusetts Inc for work performed outside of the current study and has received funding from the University of Michigan Institute for Healthcare Policy and Innovation Clinician Scholars Program and the National Institutes of Health Division of Loan Repayment. Justin B. Dimick is a cofounder of ArborMetrix Inc, a company that makes software for profiling hospital quality and efficiency, and has received grant funding from the National Institutes of Health (R01AG039434). Hari Nathan has received funding from the Agency for Healthcare Research and Quality (K08HS024763) and the National Institutes of Health (R01AG039434).

AUTHOR CONTRIBUTIONS

All authors have made substantial contributions to the conception or design of the work or the acquisition, analysis, or interpretation of the data. All authors have contributed substantially to drafting the work or revising it critically for important intellectual content. All authors have given final approval of the version to be published.

REFERENCES

1. American Hospital Association. TrendWatch Chartbook 2019. Accessed February 11, 2020. https://www.aha.org/guidesreports/2019-10-25-trendwatch-chartbook-2019
2. Comparative Health System Performance Initiative. Snapshot of U.S Health Systems, 2016. Accessed February 11, 2020. https://www.ahrq.gov/sites/default/files/wysiwyg/snapshot-of-us-health-systems-2016v2.pdf
3. Dafny LS, Lee TH. The good merger. *N Engl J Med*. 2015;372:2077-2079. doi:10.1056/NEJMp1502338
4. Schmitt M. Do hospital mergers reduce costs? *J Health Econ*. 2017;52:74-94. doi:10.1016/j.jhealeco.2017.01.007
5. Dafny L, Ho K, Lee RS. The price effects of cross-market mergers: theory and evidence from the hospital industry. *RAND J Econ*. 2019;50:286-325. doi:10.1111/1756-2171.12270
6. Cooper Z, Craig SV, Gaynor M, Van Reenen J. The price ain't right? Hospital prices and health spending on the privately insured. *Q J Econ*. 2019;134:51-107. doi:10.1093/qje/qjy020
7. White C, Whaley CM; RAND Corporation. Prices paid to hospitals by private health plans are high relative to Medicare and vary widely: findings from an employer-led transparency initiative. Published 2019. Accessed April 16, 2020. https://www.rand.org/pubs/research_reports/RR3033.html
8. Chhabra KR, Sheetz KH, Regenbogen SE, Dimick JB, Nathan H. Wide variation in surgical spending within hospital systems: a missed opportunity for bundled payment success. *Ann Surg*. Published online December 10, 2019. doi:10.1097/SLA.0000000000003741
9. Sheetz KH, Chhabra K, Nathan H, Dimick JB. The quality of surgical care at hospitals associated with America's highest-rated medical centers. *Ann Surg*. 2020;271:862-867. doi:10.1097/SLA.0000000000003195
10. Ibrahim AM, Dimick JB. Redesigning the delivery of specialty care within newly formed hospital networks. *NEJM Catal*. Published online March 30, 2017. doi:10.1056/CAT.17.0503
11. Dummit LA, Kahvecioglu D, Marrufo G, et al. Association between hospital participation in a Medicare bundled payment initiative and payments and quality outcomes for lower extremity joint replacement episodes. *JAMA*. 2016;316:1267-1278. doi:10.1001/jama.2016.12717
12. Austin AM, Gottlieb DJ, Carmichael DQ, et al. Technical Report: A Standardized Method for Adjusting Medicare Expenditures for Regional Differences in Prices. The Dartmouth Institute for Health Policy and Clinical Practice; 2018.
13. Birkmeyer JD, Gust C, Dimick JB, Birkmeyer NJO, Skinner JS. Hospital quality and the cost of inpatient surgery in the United States. *Ann Surg*. 2012;255:1-5. doi:10.1097/SLA.0b013e3182402c17
14. Birkmeyer JD, Gust C, Baser O, Dimick JB, Sutherland JM, Skinner JS. Medicare payments for common inpatient procedures: implications for episode-based payment bundling. *Health Serv Res*. 2010;45(6 pt 1):1783-1795. doi:10.1111/j.1475-6773.2010.01150.x
15. Pradarelli JC, Scally CP, Nathan H, Thumma JR, Dimick JB. Hospital teaching status and Medicare expenditures for complex surgery. *Ann Surg*. 2017;265:502-513. doi:10.1097/SLA.0000000000001706

16. Sheetz KH, Dimick JB, Nathan H. Centralization of high-risk cancer surgery within existing hospital systems. *J Clin Oncol*. 2019;37:3234-3242. doi:10.1200/JCO.18.02035
17. Iezzoni L. Risk Adjustment for Measuring Health Care Outcomes. 4th ed. Health Administration Press; 2012.
18. Hofer TP, Hayward RA, Greenfield S, Wagner EH, Kaplan SH, Manning WG. The unreliability of individual physician "report cards" for assessing the costs and quality of care of a chronic disease. *JAMA*. 1999;281:2098-2105. doi:10.1001/jama.281.22.2098
19. Dimick JB, Staiger DO, Birkmeyer JD. Ranking hospitals on surgical mortality: the importance of reliability adjustment. *Health Serv Res*. 2010;45(6 pt 1):1614-1629. doi:10.1111/j.1475-6773.2010.01158.x
20. Grenda TR, Krell RW, Dimick JB. Reliability of hospital cost profiles in inpatient surgery. *Surgery*. 2016;159:375-380. doi:10.1016/j.surg.2015.06.043
21. Nathan H, Atoria CL, Bach PB, Elkin EB. Hospital volume, complications, and cost of cancer surgery in the elderly. *J Clin Oncol*. 2015;33:107-114. doi:10.1200/JCO.2014.57.7155
22. Short MN, Aloia TA, Ho V. The influence of complications on the costs of complex cancer surgery. *Cancer*. 2014;120:1035-1041. doi:10.1002/cncr.28527
23. Short MN, Ho V, Aloia TA. Impact of processes of care aimed at complication reduction on the cost of complex cancer surgery. *J Surg Oncol*. 2015;112:610-615. doi:10.1002/jso.24053
24. Navathe AS, Emanuel EJ, Venkataramani AS, et al. Spending and quality after three years of Medicare's voluntary bundled payment for joint replacement surgery. *Health Aff (Millwood)*. 2020;39:58-66. doi:10.1377/hlthaff.2019.00466
25. Agarwal R, Liao JM, Gupta A, Navathe AS. The impact of bundled payment on health care spending, utilization, and quality: a systematic review. *Health Aff (Millwood)*. 2020;39:50-57. doi:10.1377/hlthaff.2019.00784
26. Birkmeyer JD, Warshaw AL, Finlayson SR, Grove MR, Tosteson AN. Relationship between hospital volume and late survival after pancreaticoduodenectomy. *Surgery*. 1999;126:178-183.
27. Finks JF, Osborne NH, Birkmeyer JD. Trends in hospital volume and operative mortality for high-risk surgery. *N Engl J Med*. 2011;364:2128-2137. doi:10.1056/NEJMsa1010705
28. Birkmeyer JD, Siewers AE, Finlayson EV, et al. Hospital volume and surgical mortality in the United States. *N Engl J Med*. 2002;346:1128-1137. doi:10.1056/NEJMsa012337
29. Sheetz KH, Chhabra KR, Smith ME, Dimick JB, Nathan H. Association of discretionary hospital volume standards for high-risk cancer surgery with patient outcomes and access, 2005-2016. *JAMA Surg*. 2019;154:1005-1012. doi:10.1001/jamasurg.2019.3017
30. Chhabra KR, Dimick JB. Hospital networks and value-based payment: fertile ground for regionalizing high-risk surgery. *JAMA*. 2015;314:1335-1336. doi:10.1001/jama.2015.9803
31. Nelson G, Kiyang LN, Crumley ET, et al. Implementation of Enhanced Recovery After Surgery (ERAS) across a provincial healthcare system: the ERAS Alberta Colorectal Surgery Experience. *World J Surg*. 2016;40:1092-1103. doi:10.1007/s00268-016-3472-7
32. Ljungqvist O, Scott M, Fearon KC. Enhanced recovery after surgery: a review. *JAMA Surg*. 2017;152:292-298. doi:10.1001/jamasurg.2016.4952
33. Tsai TC, Greaves F, Zheng J, Orav EJ, Zinner MJ, Jha AK. Better patient care at high-quality hospitals may save Medicare money and bolster episode-based payment models. *Health Aff (Millwood)*. 2016;35:1681-1689. doi:10.1377/hlthaff.2016.0361
34. Balentine CJ, Richardson PA, Mason MC, Naik AD, Berger DH, Anaya DA. Postacute care and recovery after cancer surgery: still a long way to go. *Ann Surg*. 2017;265:993-999. doi:10.1097/SLA.0000000000001758