Variation in payments for spine surgery episodes of care: implications for episode-based bundled payment

Elyne N. Kahn, MD, MPH,1,2 Chandy Ellimootttil, MD, MS,1,3 James M. Dupree, MD, MPH,1,3 Paul Park, MD,1,2 and Andrew M. Ryan, PhD1,4

1Institute for Healthcare Policy and Innovation, 2Department of Neurosurgery, 3Department of Urology, and 4Department of Health Management and Policy/School of Public Health, University of Michigan, Ann Arbor, Michigan

OBJECTIVE Spine surgery is expensive and marked by high variation across regions and providers. Bundled payments have potential to reduce unwarranted spending associated with spine surgery. This study is a cross-sectional analysis of commercial and Medicare claims data from January 2012 through March 2015 in the state of Michigan. The objective was to quantify variation in payments for spine surgery in adult patients, document sources of variation, and determine influence of patient-level, surgeon-level, and hospital-level factors.

METHODS Hierarchical regression models were used to analyze contributions of patient-level covariates and influence of individual surgeons and hospitals. The primary outcome was price-standardized 90-day episode payments. Intraclass correlation coefficients—measures of variability accounted for by each level of a hierarchical model—were used to quantify sources of spending variation.

RESULTS The authors analyzed 17,436 spine surgery episodes performed by 195 surgeons at 50 hospitals. Mean price-standardized 90-day episode payments in the highest spending quintile exceeded mean payments for episodes in the lowest cost quintile by $42,953 (p < 0.001). Facility payments for index admission and post-discharge payments were the greatest contributors to overall variation: 39.4% and 32.5%, respectively. After accounting for patient-level covariates, the remaining hospital-level and surgeon-level effects accounted for 2.0% (95% CI 1.1%–3.8%) and 4.0% (95% CI 2.9%–5.6%) of total variation, respectively.

CONCLUSIONS Significant variation exists in total episode payments for spine surgery, driven mostly by variation in post-discharge and facility payments. Hospital and surgeon effects account for relatively little of the observed variation.

https://thejns.org/doi/abs/10.3171/2017.12.SPINE17674

KEYWORDS spine surgery costs; bundled payments; payment models; episode-based payment; Medicare

The Centers for Medicare and Medicaid Services (CMS) is moving aggressively toward alternative payment models. Bundled payment—in which providers (hospitals or physician groups) receive a single payment for providing all care to a patient during a defined episode—is a key component of the alternative payment model reform effort. CMS has implemented the Comprehensive Care for Joint Replacement Model, a bundled payment program for Medicare patients undergoing hip and knee replacement surgery at 800 hospitals in 67 metropolitan statistical areas. Proponents of bundled payment believe that it will incentivize providers to provide higher-quality care while reducing costs. Spine surgery is a prime candidate for bundled payment reform efforts. Spending on spine surgery has increased dramatically in recent decades. Variation in spending patterns—driven in part by hospital and surgeon preferences—suggests that a substantial portion of the total cost of spine surgery care is unwarranted. The clinical nuances of spine surgery, however, present challenges. Unlike joint replacement surgery for osteoarthritis, spine surgery includes multiple types of procedures of varying intensities for a variety of indications. To achieve a good patient outcome, minimally invasive laminotomy for a focal disc herniation, for instance, requires vastly different resources than a multilevel fusion with osteotomies for symptomatic degenerative spinal deformity. Further complicating the picture is a paucity of clear consensus guidelines and appropriateness criteria associated with spine surgery. These features may contrib-
ute to wide variation in resource utilization among spine surgery patients, in both the acute and post-acute care contexts.

Understanding the sources of spending variation within spine surgery episodes is crucial for understanding the potential to limit spending variation and for predicting the success of bundled payment reforms. We explore this topic in the context of comprehensive multiyear commercial and Medicare spending data on spine surgery in the state of Michigan.

Methods

Data

We retrieved all data from the Michigan Value Collaborative (MVC). MVC is a collaboration between Blue Cross Blue Shield of Michigan and 75 hospitals in the state. The MVC data set includes both commercial and Medicare claims data for inpatient admissions at each of the participating hospitals. 10,11

Study Population

The study population consisted of adult patients (≥ 18 years) in the state of Michigan who underwent a spine-based surgical procedure at any one of the MVC participating hospitals from January 2, 2012, through September 29, 2015. We included patients whose procedures fell under the 7 most common spine-related diagnosis-related group (DRG) codes and whose indications for surgery fell under the 10 most common spine-related ICD-9 codes. (See Appendix, eTable 1.) We excluded outpatient-based procedures (n = 5,766) and episodes for which the surgeon’s national provider identifier number was missing (n = 1,645). If a surgeon practiced at multiple hospitals, we included only those episodes at the hospital at which that surgeon practiced most frequently, and we excluded all others (n = 2,113). We excluded episodes associated with hospitals and surgeons with fewer than 10 episodes each (n = 299).

Outcomes

The primary outcome was price-standardized total episode payment for each 90-day episode. Price standardization accounts for contractual differences in negotiated rates between hospitals in the data set so that payments for identical services are equal across hospitals. Total episode payments consist of inpatient facility payments (including for index hospitalization), facility payments associated with rehospitalizations, post-discharge payments, and professional payments, including those for surgical procedures, office visits, consultations, ancillary testing, and outpatient rehabilitation. This is similar to Medicare’s 90-day Global Surgery Package. One distinction is that preoperative visits on the day before surgery are not included in the 90-day episodes analyzed in this study, whereas they are included in Medicare’s 90-day definition. To understand the overall variation in episode payments, we categorized episodes into quintiles according to total episode payments.

Analysis

We evaluated the contributions of individual components (facility payments, professional payments, post-discharge payments, and readmission payments) to the total variation. We then calculated the extent to which individual components of post-discharge spending (inpatient rehabilitation, outpatient rehabilitation, skilled nursing, home health, and other post-discharge payments) contributed to variation in total post-discharge payments.

We performed several analyses to identify the influence of patient-level, surgeon-level, and hospital-level factors on spending. First, to assess surgical practice patterns, we calculated the share of each surgeon’s procedures that were cervical fusions (DRGs 471, 472, and 473), thoracolumbar fusions (DRGs 459 and 460), and non-fusion procedures (DRGs 490 and 491). We then performed a multilevel (hierarchical) mixed-effects linear regression in which each observation (patient care episode) was nested within a group defined by individual surgeons and each surgeon was nested within a group defined by individual hospitals. The hierarchical model allowed us to account for and analyze the contributions of patient-level covariates and the influence of individual surgeons and hospitals. 6,13 The dependent variable for the model was price-standardized total payment. Included in the model were independent variables from 4 different categories of patient-level characteristics. The first category was demographics, including age, sex, and insurance status (Medicare vs commercial insurance). The second category was comorbidities, defined in this cohort by Medicare’s 70 hierarchical condition categories. The third category was diagnosis (indication for surgery), defined in this cohort by ICD-9 code. We included this category because the diagnosis codes provide insight into the specific nature and severity of a patient’s condition—for instance, whether a patient had spinal cord dysfunction (myelopathy) preoperatively—and these conditions would be expected to influence patient management for reasons that are distinct from surgeon or hospital preferences. The fourth category was the procedure type and case mix, defined in this cohort by DRG code. Group-level variables were surgeon (by national provider identifier) and hospital (by facility identification number). To quantify the contributions of hospital effects and surgeon effects, we computed intraclass correlation coefficients for hospitals and surgeons. The intraclass correlation coefficient is a means of describing the variability accounted for by each level of the hierarchical model. 6,13

We performed several sensitivity analyses to confirm the stability of the results. One analysis eliminated surgeons from the regression model to create a 2-level model in which individual episodes were nested within each hospital. We then analyzed index facility payments alone and post-discharge payments alone. To account for the possibility that surgeon preferences—rather than patient factors alone—might influence choice of procedure, a third sensitivity analysis omitted procedure type (DRG codes) from the regression model. Finally, in order to determine whether our findings apply to specific procedure types, we repeated the total episode payment analyses separately for non-fusion procedures, thoracolumbar fusion procedures, and cervical fusion procedures.

All analyses were performed using statistical software (Stata 14/MP, StataCorp) at the 5% significance level. This
study was deemed exempt by the University of Michigan Medical School Institutional Review Board.

Results

The final data set included 17,436 spine surgical episodes performed by 195 surgeons at 50 hospitals. The median surgeon and hospital volumes were, respectively, 68 (interquartile range [IQR] 31–132) and 292 (IQR 102–458) (Table 1).

Mean price-standardized 90-day episode payments in the highest spending quintile exceeded mean payments for episodes in the lowest cost quintile by $42,953 (p < 0.001) (Table 2). When evaluating the individual components of each 90-day episode, facility payments for the index admission and post-discharge payments were the greatest contributors to total variation (39.4% and 32.5%, respectively; Table 2 and Fig. 1).

After categorizing episodes into quintiles according to post-discharge payments, we found that average payments for each component of post-discharge care varied significantly between the highest cost quintile and the lowest cost quintile (see Appendix, eTable 2). The components of care with the greatest contributions to the total variation in post-discharge payments were skilled nursing facilities (42.7%), inpatient rehabilitation (29.0%), and home health (21.9%).

Across surgeons, we observed substantial variation with respect to procedure mix. Per surgeon, the mean rate of cervical fusions was 36.8% (IQR 27.5%–46.2%), the mean rate of thoracolumbar fusions was 29.7% (IQR 17.1%–49.7%), and the mean rate of non-fusion procedures was 30.4% (IQR 13.1%–44.0%). In this study cohort the mean episode payment was $25,662 (SD $14,627) for cervical fusions, $39,317 (SD $15,075) for thoracolumbar fusions, and $16,152 (SD $13,382) for non-fusion procedures.

In the regression analysis, patient-level characteristics in each category (demographics, comorbidities, indication for procedure, and procedure type) were significantly associated with variation in total episode payments (see Appendix, eTable 3a). After accounting for patient-level covariates, the intraclass correlation coefficient was 0.020 for hospitals (95% CI 0.011–0.038) and 0.040 for surgeons (95% CI 0.029–0.056), suggesting that hospital-level effects and surgeon-level effects accounted for, respectively, only 2.0% and 4.0% of the total variation in episode payments in this cohort (Table 3, Model 1).

Sensitivity Analyses

Our first sensitivity analysis eliminated surgeons from the regression model. In this model—which did not control for the influence of individual surgeons on overall variation—the intraclass correlation coefficient suggested that hospital-level effects accounted for 2.6% (95% CI 1.6%–4.1%) of the variation in total episode payments (Table 3, Model 2).

Analyses of index facility payments alone and post-discharge payments alone demonstrated that—as with total episode payments—a number of patient-level characteristics were significantly associated with payment variation (Appendix, eTables 3b and 3c). After accounting for patient-level covariates, the models suggested that hospital-level effects and surgeon-level effects accounted for, respectively, only 0.7% (95% CI 0.3%–1.6%) and 1.6% (95% CI 1.0%–2.4%) of the total variation in index facility payments and only 2.0% (95% CI 1.1%–3.7%) and 3.3% (95% CI 2.2%–4.7%) of the total variation in post-discharge payments in this cohort (Table 3, Model 1).

The next sensitivity analysis omitted procedure types (DRG codes) as covariates from the regression model to

TABLE 1. Characteristics of study cohort

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient characteristics (n = 17,436)</td>
<td></td>
</tr>
<tr>
<td>Age in yrs, mean (SD)</td>
<td>61.0 (13.6)</td>
</tr>
<tr>
<td>Sex, n (%)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>8,648 (49.6)</td>
</tr>
<tr>
<td>Male</td>
<td>8,788 (50.4)</td>
</tr>
<tr>
<td>Insurance type, n (%)</td>
<td></td>
</tr>
<tr>
<td>Medicare</td>
<td>10,151 (58.2)</td>
</tr>
<tr>
<td>Commercial</td>
<td>7,285 (41.8)</td>
</tr>
<tr>
<td>Procedure type, n (%)</td>
<td></td>
</tr>
<tr>
<td>Cervical fusions</td>
<td>6,537 (37.5)</td>
</tr>
<tr>
<td>Thoracolumbar fusions</td>
<td>5,480 (31.4)</td>
</tr>
<tr>
<td>Non-fusion procedures</td>
<td>5,419 (31.1)</td>
</tr>
<tr>
<td>Surgeon characteristics (n = 195)</td>
<td></td>
</tr>
<tr>
<td>No. of procedures, median (IQR)</td>
<td>68 (31–132)</td>
</tr>
<tr>
<td>Hospital characteristics (n = 50)</td>
<td></td>
</tr>
<tr>
<td>No. of procedures, median (IQR)</td>
<td>292 (102–458)</td>
</tr>
</tbody>
</table>

TABLE 2. Payment variation by category (n = 17,436)

<table>
<thead>
<tr>
<th>Payment Variable</th>
<th>Overall Mean</th>
<th>1st Quintile Mean</th>
<th>5th Quintile Mean</th>
<th>Difference btwn 1st &amp; 5th Quintiles</th>
<th>Percentage of Total Variation*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total episode payment</td>
<td>26,998</td>
<td>9,814</td>
<td>52,767</td>
<td>42,953</td>
<td>100.0%</td>
</tr>
<tr>
<td>Index facility payment</td>
<td>16,650</td>
<td>7,062</td>
<td>24,000</td>
<td>16,938</td>
<td>39.4%</td>
</tr>
<tr>
<td>Post-discharge payment</td>
<td>3,908</td>
<td>605</td>
<td>14,561</td>
<td>13,956</td>
<td>32.5%</td>
</tr>
<tr>
<td>Readmission payment</td>
<td>1,328</td>
<td>4</td>
<td>5,668</td>
<td>5,664</td>
<td>13.2%</td>
</tr>
<tr>
<td>Professional payment</td>
<td>5,044</td>
<td>2,252</td>
<td>8,195</td>
<td>5,943</td>
<td>13.8%</td>
</tr>
</tbody>
</table>

Mean values and differences are in US dollars and are rounded to the nearest dollar.

* Defined as the ratio of difference between the 1st and 5th quintiles for each variable to the difference between the 1st and 5th quintiles for total episode payment. For instance, the percentage of total variation attributable to index facility payment was calculated as $16,938/$42,953 = 39.4%.
account for the possibility that surgeon preferences might influence choice of procedure. In this case, the model suggested that hospital-level effects and surgeon-level effects accounted for, respectively, 3.9% (95% CI 2.0%–7.5%) and 10.4% (95% CI 8.0%–13.4%) of the total variation in episode payments in this cohort (Table 3, Model 3).

Finally, we repeated the analysis for each of three main categories of procedure: non-fusion procedures, thoracolumbar fusions, and cervical fusions. In each case, post-discharge payments represented the greatest contributor to total variation (Appendix, eTables 4a, 5a, and 6a). Intraclass correlation coefficients remained low, ranging from 0.011 to 0.032 for hospitals and from 0.033 to 0.073 for surgeons in the analysis of total episode payments (Appendix, eTables 4b, 5b, and 6b).

Discussion

A large cohort of patients receiving spine surgery in Michigan demonstrated significant variation in total episode payments for spine surgery. Facility payments and post-discharge payments were the largest component contributors to total payment variation. Variation in total episode payments was attributable largely to patient-level characteristics. Hospital and surgeon effects, on the other hand, account for relatively little of the observed variation.

In post-discharge payments specifically, variation was driven mostly by payments for skilled nursing facilities, inpatient rehabilitation, and home health services. The analyses of index facility payments and post-discharge payments again demonstrated that variation is accounted for largely by patient-level characteristics, with minor contributions from the effects of individual hospitals and surgeons.

These findings challenge the concept that provider and hospital preferences substantially affect payments for individual episodes of care for spine surgery. Collectively, the findings suggest that patient-level characteristics are important predictors of episode costs.

The findings of this study are consistent with previous investigators’ attempts to describe variation in payments associated with spine surgery. This report is the first, however, to quantify the contributions of patient-level versus provider- and hospital-level effects on this variation and the first to examine the detailed contributions of various types of post-discharge payments.

### Table 3. Surgeon and hospital contributions to variation in payments (n = 17,436)

<table>
<thead>
<tr>
<th>Model No.</th>
<th>Hospital Effect</th>
<th>Surgeon Effect</th>
<th>Procedure Types Omitted</th>
<th>Total Episode Payments Hospitals</th>
<th>Total Episode Payments Surgeons</th>
<th>Index Admission Facility Payments Hospitals</th>
<th>Index Admission Facility Payments Surgeons</th>
<th>Post-Discharge Payments Hospitals</th>
<th>Post-Discharge Payments Surgeons</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>0.020</td>
<td>0.040</td>
<td>0.007</td>
<td>0.016</td>
<td>0.020</td>
<td>0.033</td>
</tr>
<tr>
<td>2</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>0.026</td>
<td>NA</td>
<td>0.010</td>
<td>NA</td>
<td>0.026</td>
<td>NA</td>
</tr>
<tr>
<td>3</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>0.039</td>
<td>0.104</td>
<td>0.054</td>
<td>0.172</td>
<td>0.020</td>
<td>0.035</td>
</tr>
</tbody>
</table>

NA = not applicable.
This study has several limitations. First, the analysis included only episodes of care in the state of Michigan and only patients insured by Medicare or a single large private insurance company. Generalizability to other geographic regions and to other payer mixes may be limited. Second, no information was available about the functional status or postoperative outcomes of the patients in this cohort. Third, the study does not evaluate the appropriateness of the surgical procedures performed or the association of payments with patient outcomes.

Despite these limitations, the results should help to inform stakeholders as they seek to address unwarranted variation in costs of care. Relevant to policymakers and payers, this research suggests that hospital effects and provider preferences may not play a substantial role in variation in payments. These findings are especially relevant to the design of alternative payment models that are meant to limit unwarranted variation in costs. Bundled payment, for instance, seeks to limit variation and reduce costs by uncoupling payments from individual services and thereby reducing provider incentive to provide unnecessary care.1 Our results suggest that the opportunities for such cost savings may be limited, given the limited contribution of provider- and hospital-level effects to variation in spine surgery payments. In the context of bundled payment, surgeons and hospitals may find it challenging to limit variation when the factors underlying that variation are not simply a function of their own preferences. That said, tools are available to help providers identify and manage unwarranted variation, including prospective databases to collect and analyze relevant practice data, consensus guidelines to help ensure that care provided is appropriate to a patient’s condition, and patient education to encourage informed participation in the decision-making process for elective care. Finally, thoughtfully designed alternative payment models have the potential to encourage practice redesign to better manage variation.21 Our results particularly underscore the importance of careful risk adjustment in the context of bundled payment. Also relevant to stakeholders with an interest in containing costs of care, variation in episode payments was most strongly associated with the use of skilled nursing facilities, inpatient rehabilitation, and home health services. These services may therefore serve as more appropriate targets for cost containment compared with other components of post-discharge care, such as outpatient rehabilitation.

Conclusions

Future research should focus on further understanding the contributors to variation that remain unexplained. These efforts will be aided by careful examination of patient characteristics included in clinical registry data. Additional research should focus on enhancing risk adjustment and payment modifiers in the context of bundled payments and on defining guidelines and appropriateness criteria for spine surgery and associated post-discharge care. Ultimately, an improved understanding of which factors contribute to cost variation will help us to achieve the dual goal of lower costs and better patient outcomes.

Acknowledgments

We would like to acknowledge the data support of John Syrjamaki. Support for the Michigan Value Collaborative is provided by Blue Cross Blue Shield of Michigan as part of the Blue Cross Blue Shield of Michigan Value Partnerships program; however, the opinions, beliefs, and viewpoints expressed by the authors do not necessarily reflect those of Blue Cross Blue Shield of Michigan or any of its employees.

References


**Disclosures**

Drs. Dupree and Chandy Ellimoottil have grant funding from Blue Cross Blue Shield of Michigan. Dr. Park has grant funding from Pfizer Inc. and Vertex Pharmaceuticals, receives royalties from Globus Medical, Inc., and has received consulting fees from Medtronic, Globus Medical, Inc., Zimmer Biomet, and NuVasive, Inc.

**Author Contributions**

Conception and design: Kahn, Ellimoottil, Ryan. Acquisition of data: Kahn, Ellimoottil, Dupree. Analysis and interpretation of data: all authors. Drafting the article: Kahn, Ellimoottil. Critically revising the article: all authors. Reviewed submitted version of manuscript: all authors. Approved the final version of the manuscript on behalf of all authors: Kahn. Statistical analysis: Kahn, Ellimoottil, Ryan. Study supervision: Ryan.

**Supplemental Information**

Online-Only Content

Supplemental material is available with the online version of the article.


**Correspondence**

Elyne Kahn: University of Michigan Health System, Ann Arbor, MI. kahne@med.umich.edu; elyne.kahn@gmail.com.