

The Stability of Physician-Specific Episode Costs for Urologic Cancer Surgery: Implications for Urologists Under the Merit-Based Incentive Program



Deborah R. Kaye, Zaojun Ye, Jonathan Li, Lindsey A. Herrel, James M. Dupree, Chad Ellimoottil, and David C. Miller

OBJECTIVE	To evaluate the stability of physician-specific episode payments for prostatectomy, nephrectomy, and cystectomy in the context of value-based purchasing programs, such as the merit-based incentive payment system.
METHODS	We utilized Surveillance, Epidemiology and End Results-Medicare data to identify patients aged 66-99 who underwent a prostatectomy, nephrectomy, or cystectomy from 2008 to 2012. We calculated each surgeon's average 90-day episode payment by procedure. Next, we examined payment differences between the most and least expensive quartile providers. For the most expensive quartile of physicians in 2010, we examined their spending quartile in 2011. Finally, we evaluated the correlation in spending over time and across procedures.
RESULTS	We identified 14,585 patients who underwent surgery by one of 1895 unique clinicians. Differences in payments between the highest and lowest quartiles were \$5881, \$17,714, and \$40,288 for prostatectomy, nephrectomy, and cystectomy, respectively. Only 39%, 16%, and 13% of physicians that were in the highest spending quartile for prostatectomy, nephrectomy, and cystectomy in 2010 were also in the most expensive quartile in 2011. Although we observed weak correlation in year-to-year spending for prostatectomy (0.108, $P = .033$ to $.270$, $P < .001$), annual payments for nephrectomy and cystectomy were not significantly correlated. Finally, there was minimal correlation in surgeon spending across procedures.
CONCLUSION	There is wide variation in physician-specific episode payments for prostatectomy, nephrectomy, and cystectomy. However, physician spending patterns are not stable over time or across procedures, raising concerns about the ability of the cost-based measures in merit-based incentive payment system to change physician behavior and reliably distinguish those providing less efficient or lower quality care. UROLOGY 123: 114–119, 2019. © 2018 Elsevier Inc.

The Centers for Medicare and Medicaid Services (CMS) initiated the Merit-Based Incentive Program (MIPS) to increase health care quality and

decrease expenditures. MIPS was initiated on January 1, 2017. Nearly 1.5 million clinicians will take part in MIPS, receiving performance-based awards or penalties of up to 4% of Medicare Part B payments in 2019, and up to 9% in 2022. Clinicians who bill Medicare Part B > \$30,000 per year and provide care for more than 100 Medicare Part B beneficiaries are included in MIPS. Exclusion criteria include clinicians who are in non-MIPS eligible specialties, those who are in their first year of Medicare participation, and clinicians who participate in MACRA as a Qualifying Alternative Payment Model Participant or Partial Qualifying Participant. Participating clinicians include physicians, physician assistants, nurse practitioners, clinical nurse specialists, and certified registered nurse anesthetists.^{1,2} The resource use measure is one of four categories (quality, Meaningful Use of Certified Electronic Health Records, clinical practice

Financial Disclosure: David C. Miller received a grant contract from the Blue Cross Blue Shield of Michigan for the Michigan Urological Surgery Improvement Collaborative. James M. Dupree received a grant contract from the Blue Cross Blue Shield of Michigan for the Michigan Value Collaborative and the Michigan Urological Surgery Improvement Collaborative. The remaining authors declare that they have no relevant interests. Deborah R. Kaye has received grant funding from the Blue Cross Blue Shield of Michigan Foundation.

Funding: This project was supported by the National Cancer Institute (1-R01-CA-174768-01-A1 to DCM & 5-T32-CA-180984-03 to DRK).

From the Dow Division of Health Services Research, Department of Urology, University of Michigan, Ann Arbor, MI

Address correspondence to: Deborah R. Kaye, M.D., M.Sc., Dow Division for Urologic Health Services Research, University of Michigan, NCRC Building 16, Room 108E, 2800 Plymouth Road, Ann Arbor, MI 48109-2800. E-mail: kayed@med.umich.edu

Submitted: January 24, 2018, accepted (with revisions): April 27, 2018

improvement activities, and resource use) that comprise the MIPS composite score; this metric specifically evaluates clinician spending for attributed Medicare beneficiaries. The goal of the resource use measure is to increase clinician awareness and accountability for health care expenditures.³ By 2021, this measure will account for 30% of the composite score. From a measurement perspective, CMS plans to evaluate the total health care expenditures for all patients or episodes of care attributed to an individual clinician.

Although the implementation details are still being finalized, the potential impact of the resource use measure is predicated on two key principles. The first principle, established in the literature for some diagnoses and procedures, is that wide variation exists in physician-specific episode payments.⁴⁻⁹ The second principle, not yet tested empirically, is that the care provided by some physicians is consistently more expensive. In other words, there is a group of providers that are at the high end of the overall spending measure both year-over-year and across many procedures and/or diagnoses. In fact, the impact of any resource use measure will arguably be greatest if these clinicians consistently have a large and consistent incentive to change behavior in a way that reduces resource utilization. Likewise, reliable and actionable payment data indicating that physicians are consistently more expensive than their peers may help identify areas of poor quality and/or inefficient care delivery.^{10,11}

In contrast, the absence of stability in physician spending, defined in this manuscript, as similarities in physician payments over time and across procedures, has two important implications for the potential impact of the resource use measure. First, if physician spending signatures are not stable, the financial incentives associated with a resource use measure will be much less apparent. For instance, a clinician who is a high spender for one procedure, but not for another, will have a spending pattern that is similar to the average physician across procedures, and thus little imperative to change behavior. Second, although there are other measures of quality in the MIPS program, absence of stability in the resource use measure will greatly limit its ability to provide complementary information (to the MIPS quality metrics) that identifies, and motivates change, among physicians who are expensive due to concerns such as complications or decisions around the use of discretionary services.

In this context, evaluating surgical episode spending is a useful potential litmus test for measuring stability in physician-specific episode payments. Surgical admissions comprise one-third of all inpatient admissions and accounted for the greatest increase in acute inpatient spending in 2014.¹² Moreover, because a designated clinician performs a procedure, attribution of surgical episodes, and therefore measuring physician resource utilization, may be more straightforward than attribution of a nonprocedure-based inpatient hospitalization.

Accordingly, we used linked Surveillance, Epidemiology and End Results (SEER)-Medicare data to evaluate physician-specific episode payments for prostatectomy,

nephrectomy, and cystectomy. We further examined whether procedure-specific spending was stable over time and if spending for one genitourinary cancer procedure was associated with high spending for the other procedures. We hypothesize that, while variation in spending exists for prostatectomy, nephrectomy, and cystectomy, there is little stability in patterns of clinician spending over time and/or across procedures. By virtue of this approach, our findings will inform physicians, payers, and policymakers on variation in physician-specific episode payments and the ability of the resource use measure to influence physician behavior, consistently differentiate physician spending, and provide reliable and actionable information to decrease expenditures while maintaining or improving quality.

METHODS

Data

We used data from SEER registries linked with Medicare claims data from 2008 through 2012. SEER-Medicare is a patient-level dataset that links Medicare claims with clinical, demographic, and outcome data from SEER cancer registries covering 17 major regions across the United States. We used the following Medicare claims files for our analyses: 100% Medicare Provider and Analysis Review File, carrier, outpatient, home health and hospice.

Study Population

Our study cohort included patients aged 66-99 years who underwent a prostatectomy, nephrectomy, or cystectomy for cancer. We identified these patients using International Classification of Diseases, Ninth Edition (ICD-9) diagnosis and procedure codes and current procedural terminology codes (Appendix Table 1). We excluded cases that were missing an ICD-9 diagnosis or procedure code for the cancer of interest, a current procedural terminology code for the surgery of interest, and/or a cancer diagnosis listed in the SEER file. We further excluded patients who underwent multiple surgeries for different cancers, those that did not have continuous Medicare Parts A and B enrollment and/or those who participated in Medicare Health Maintenance Organizations during the study interval.

Surgeon Attribution

In order to evaluate surgeon episode payments, we attributed each episode to a single surgeon. To do this, we first identified the national provider identifier associated with the surgery claim. If more than one surgeon was associated with the claim, we attributed the episode to the surgeon with the HCFA specialty code for urology (34). If a urologist was not listed, the claim was attributed to 02 (general surgery) or 91 (surgical oncology). If more than one provider with the appropriate HCFA specialty code was associated with the index surgery, the episode was attributed to the provider with the greatest payment amount.

Estimating 90-Day Episode Payments

We first identified the index hospitalization for each patient meeting our inclusion criteria. We then aggregated all payments for claims that occurred 2 days prior to the initial hospital admission through 90 days after discharge. The 90-day episode is consistent with the global surgery period and other episode-based

payment programs.¹³ To account for differences in Medicare payments based on disproportionate share payments, teaching status, and geography, we price standardized all payments using previously described methods, which were adapted from the Dartmouth Institute for Health Policy and Clinical Practice and the Medicare Payment Advisory Commission.¹⁴ In addition, we used established methods to decrease the effect of outliers on spending, by excluding patients with episode payments below the first and above the 99th percentile. This approach is consistent with CMS's methods in performance programs, including the physician quality and resource use reports and the calculation for the Medicare spending per beneficiary measure.¹⁵⁻¹⁷

Examining Provider-Level Variation in 90-Day Episode Payments

To evaluate provider-level variation in 90-day episode payments, we calculated the average episode payment per surgeon for each procedure. Similar to the proposed MIPS methodology which ranks physicians according to spending, we sorted each physician's average episode payment into quartiles. We next evaluated the mean difference between the highest and lowest quartile physicians for each type of cancer surgery.

To assess the stability in spending patterns over time, we evaluated whether the top spenders in 1 year were also top spenders in the subsequent year. Specifically, we examined the spending quartile in 2011 for surgeons in the most expensive quartile in 2010. We further evaluated the correlation in provider spending rank for each year by procedure for 2008 to 2012 using the spearman rank correlation coefficient.

Finally, to evaluate the relationship between surgeon-specific expenditures across different procedures, we ranked each surgeon by his or her mean 90-day episode payment per procedure and evaluated the spearman rank correlation coefficient for surgeon rank across procedures (ie, prostatectomy, nephrectomy, or cystectomy). All analyses were performed using SAS v 9.4 (Cary, NC) and STATA (College Station, TX). The University of Michigan's Institutional Review Board deemed this study exempt from review.

RESULTS

We identified 14,585 patients who underwent a major urologic cancer surgery by one of 1895 surgeons from 2008 through 2012. The distribution of cancer sites was as follows: 60% prostate (8673 patients and 1250 surgeons), 28% kidney (4111 patients and 1277 surgeons), and 12% bladder (1801 patients and 673 surgeons).

We observed wide variation in episode payments for each type of cancer surgery. Differences in mean payments between the highest and lowest surgeon quartiles were \$5881, \$17,714, and \$40,288 for prostatectomy, nephrectomy, and cystectomy, respectively, reflecting relative differences of 40%, 62%, and 67% of the mean episode payments of the high quartile surgeons (Table 1).

Surgeon-spending was only weakly correlated from 1 year to the next. Among the top quartile spenders in 2010, only 39%, 16%, and 13% were also in the top spending quartile in 2011 for prostate, kidney, and bladder cancer, respectively (Fig. 1). There was minimal correlation in year-over-year spending for prostate cancer, with Spearman rank correlation coefficients ranging from 0.108 (2008 to 2011, $P = .033$) to 0.270 (2009 to 2011, $P < .001$). Likewise, correlations between year-over-year payments were weak for surgeons performing cystectomy and nephrectomy, and none of these achieved statistical significance (Table 2).

Finally, high spending for surgical episodes with one urologic cancer was poorly correlated with high spending for another urologic cancer. Spearman rank correlation coefficients for surgeon-specific episode spending were 0.028 ($P = .447$) for prostate vs kidney, 0.068 ($P = .157$) for prostate vs bladder, and 0.098 ($P = .037$) for kidney vs bladder (Table 2).

DISCUSSION

We evaluated 90-day episode payments for Medicare beneficiaries undergoing major surgery for prostate, kidney, or bladder cancer to determine both variations in spending across providers, and the stability in spending rank over time and across procedures. We identified wide variation in physician-specific episode payments. In fact, for cystectomy, high spenders had average episode costs nearly 300% of those for less expensive surgeons. At the same time, however, there was little correlation between surgeon-specific episode payments over time and across procedures. Collectively, these findings suggest that, while providing physicians with payment data makes sense from the perspective of highlighting variation in resource use, metrics based on episode payments alone may be limited in their ability to change physician behavior and reliably distinguish those providing less efficient or lower quality care.

Our observation of wide variation in episode payments is consistent with, and expands upon, prior analyses demonstrating significant differences in physician-specific spending.⁴⁻⁹ An analysis of variation in Medicare Part B spending for patients hospitalized for nonelective conditions determined that payments varied more across physicians than hospitals, and that higher payments were not associated with lower 30-day readmissions or mortality rates.⁵ In an analysis of variation in 90-day payments for radical cystectomy using a nationally representative all-payer database of inpatient claims, the authors likewise found wide physician-specific variation in payments and that postoperative complications contributed the greatest to variation, followed by patient factors, including Charlson co-morbidity index, insurance type, and race.⁶

Table 1. Differences in mean 90-day surgical episode payments per physician, by spending rank quartile for prostate, kidney, and bladder cancer

Procedure	Low Payment Quartile	High Payment Quartile	Difference High and Low Payment	Difference as a % of High Payment Quartile
Prostatectomy	\$8,829	\$14,710	\$5,881	40
Nephrectomy	\$11,013	\$28,727	\$17,714	62
Cystectomy	\$19,753	\$60,041	\$40,288	67

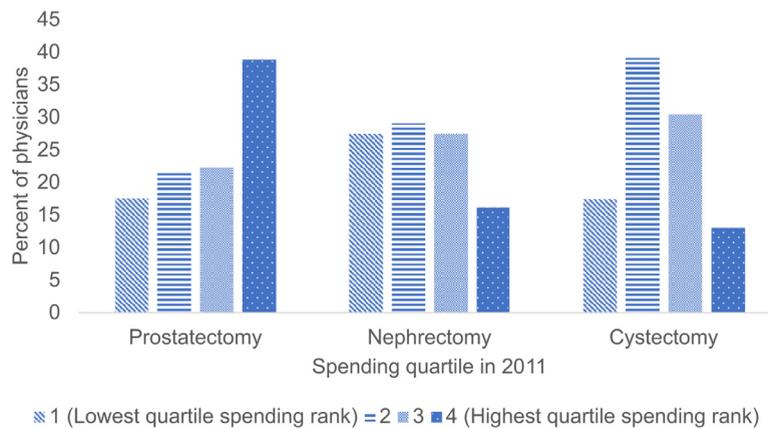


Figure 1. 2011 spending quartile for physicians that were in the top (ie, most expensive) quartile of spending in 2010. (Color version available online.)

Moreover, our finding of little correlation between physician-specific spending over time and across procedures is also consistent with other reports calling into question the reliability of physician-specific measures of quality and cost.¹⁸⁻²⁰ In an evaluation of physician spending for homogenous episodes of care for different clinical conditions, the authors determined that the majority of physicians had reliability scores well below the commonly established threshold of 0.70, with the surgical subspecialties having particularly low reliability scores. In addition, when the authors categorized physicians into 2 tiers, either low cost (ie, the lowest 25%) or not low cost, using the previously calculated cost-based reliabilities, and compared the calculated assignment to physicians' actual spending categories, 50% of orthopedic surgeons and 67% of vascular surgeons were misclassified as low cost providers, when they were in fact not low cost providers.¹⁹ Although sample size plays an important role in reliability, even substantial increases in sample size did not produce

reliability results above the commonly accepted threshold for many specialties. Another analysis of surgeon-specific complications after colectomy, demonstrated that 86% of the variability in complication rates was explained by statistical noise, and that the vast number of surgeons do not perform enough colectomies to generate reliable complication rates.¹⁸

Our study has several limitations. First, we evaluate stability in physician spending for 90-day episode payments so our findings cannot be directly generalized to the MIPS resource use measure, which has yet to be finalized. We evaluated a 90-day episode window for this analysis because it is consistent with other episode-based payment programs.¹³ Second, our evaluation did not adjust for case-mix (eg, age, comorbidity) or diagnosis-related group payment differences, which may impact episode spending. Whether and how to risk-adjust performance measures remains a controversial topic, and the risk-adjustment methodology that CMS will implement for cost-based measures in MIPS has not been finalized.³ Medicare has

Table 2. Correlation in mean 90-day surgical episode payments (A) Over time and (B) by urologic procedure

(A) Over time					
Prostatectomy					
		2008	2009	2010	2011
2008		1			
2009		0.224 (<0.001)			
2010		0.234 (<0.001)	0.270 (<0.001)		
2011		0.108 (0.033)	0.165 (<0.001)	0.184 (<0.001)	1
Nephrectomy					
		2008	2009	2010	2011
2008		1			
2009		0.016 (0.850)			
2010		0.030 (0.729)	-0.038 (0.648)		
2011		0.280 (0.001)	0.081 (0.348)	-0.043 (0.615)	1
Cystectomy					
		2008	2009	2010	2011
2008		1			
2009		-0.042 (0.660)			
2010		0.122 (0.230)	-0.026 (0.802)		
2011		0.074 (0.482)	0.127 (0.224)	0.0620 (0.553)	1
(B) By procedure					
	Prostatectomy	Nephrectomy	Cystectomy		
Prostatectomy	1				
Nephrectomy	0.028 (0.447)	1			
Cystectomy	0.068 (0.157)	0.098 (0.037)	1		

risk-adjusted in certain value-based payment programs (eg, Hospital Readmission Reduction program), while not in others (eg, proposed cardiac bundle payment program).²¹ Accurate risk adjustment depends on documentation and coding accuracy; in addition, resource availability and has been found to be lower across public hospitals compared to private for-profit and nonprofit, thus raising concern about using case-mix index as a marker of patient disease severity.²² Nonetheless, we are planning separate analyses to address this relevant issue. Third, we focus on episode spending for three inpatient surgical procedures and these findings may not be generalizable to other surgical or medical episodes. However, surgical episodes account for one-third of all inpatient admissions¹² and the included procedures are all predominantly performed by one specialty type (urology), making an evaluation of surgeons across procedures feasible. Fourth, we do not evaluate quality, which may vary across surgeons and impact the differences in spending reported herein. Put another way, our analysis does not answer the question as to whether higher spending surgeons achieve better outcomes for their patients.

These limitations notwithstanding, our results have significant implications for physicians, policymakers, and payers. For physicians, our analyses highlight wide payment variation in physician-specific payments; the availability and dissemination of such data to clinicians is likely a necessary step toward improving efficiency in care delivery. At the same time, however, for policymakers and payers our analyses highlight some cautionary notes to be considered when developing payment programs based on physician-specific data. In particular, the absence of relative stability in surgeon-specific, episode-based resource use measures for individual surgeons over time and across procedures has two important implications. First, if payers are planning to leverage this measure from a quality perspective, it may be difficult to use it to identify consistent patterns of in hospital-based or postacute spending that reflect complications, use of discretionary services, or other considerations that could be actionable from a quality or payment perspective. Furthermore, an individual surgeon may have little control over some components of episode spending (eg, intensity of services provided at a post-acute care facility), further limiting the impact of the resource use measure on decreasing payments.

The second implication relates to the magnitude of the incentive required for a physician to spend the significant time, energy, and resources needed to make meaningful system-level changes to address increased resource utilization. If there really were consistently expensive surgeons across time and procedures, then the MIPS resource use measure would have the potential to provide a consistent signal and incentive to catalyze the substantial efforts required to define and decrease resource utilization. However, if each physician ends up around the average, or if the measure varies widely year-over-year, it may be difficult to

maintain the physician focus and energy required to implement some of these changes. Collectively, our findings may be viewed by many as significant limitations with the current MIPS program; the impact of these concerns will need to be examined as the program evolves. Moving forward, it will be important to evaluate 90-day episode payments for other surgical and medical conditions, assess how risk-adjustment impacts episode spending, and if there are any other measures of resource utilization that may be more reliable, beneficial, and actionable. It is only with accurate, reliable, and clinically meaningful measures that we can improve health care quality and decrease costs.

SUPPLEMENTARY DATA

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.urology.2018.04.051](https://doi.org/10.1016/j.urology.2018.04.051).

References

1. CMS. The merit-based incentive program. 2016. Available at: <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/Merit-based-Incentive-Payment-System-MIPS-Overview-slides.pdf>.
2. CMS. How is MIPS participation determined? 2017. Available at: <https://qpp.cms.gov/participation-lookup/about>.
3. CMS. Episode-based cost measure development for the quality payment program. 2016. Available at: <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/questions-to-improve-the-cost-measure-development-process.pdf>.
4. Williams SB, Amarasekera CA, Gu X, et al. Influence of surgeon and hospital volume on radical prostatectomy costs. *J Urol*. 2012;188:2198-2202.
5. Tsugawa Y, Jha AK, Newhouse JP, Zaslavsky AM, Jena AB. Variation in physician spending and association with patient outcomes. *JAMA Internal Med*. 2017;177:675-682.
6. Leow JJ, Cole AP, Seisen T, et al. Variations in the costs of radical cystectomy for bladder cancer in the USA. *Eur Urol* 2017. Advance online publication. <https://doi.org/10.1016/j.eururo.2017.07.016>.
7. Cole AP, Leow JJ, Chang SL, et al. Surgeon and hospital level variation in the costs of robot-assisted radical prostatectomy. *J Urol*. 2016;196:1090-1095.
8. Ejaz A, Gani F, Kim Y, Pawlik TM. Variation in inpatient hospital and physician payments among patients undergoing general versus orthopedic operations. *Surgery*. 2016;160:1657-1665.
9. Chotai S, Sivaganesan A, Sielatycki JA, et al. Surgeon-level variability in outcomes, cost, and comorbidity adjusted-cost for elective lumbar decompression and fusion. *Neurosurgery*. 2018 Apr;84:506-515.
10. Grenda TR, Krell RW, Dimick JB. Reliability of hospital cost profiles in inpatient surgery. *Surgery*. 2016;159:375-380.
11. Birkmeyer JD, Gust C, Dimick JB, Birkmeyer NJ, Skinner JS. Hospital quality and the cost of inpatient surgery in the United States. *Ann Surg*. 2012;255:1-5.
12. Heath Care Cost Institute. 2013 Health care cost and utilization report. 2014. Available at: <http://www.healthcostinstitute.org/files/2013%20HCCUR%2012-17-14.pdf>.
13. Iorio R, Clair AJ, Inneh IA, Slover JD, Bosco JA, Zuckerman JD. Early results of medicare's bundled payment initiative for a 90-day total joint arthroplasty episode of care. *J Arthroplasty*. 2016;31:343-350.
14. Gottlieb DJ, Zhou W, Song Y, Andrews KG, Skinner JS, Sutherland JM. Prices don't drive regional Medicare spending variations. *Health Affairs (Project Hope)*. 2010;29:537-543.

15. CMS. Medicare Spending Per Beneficiary (MSPB) Measure Overview. QualityNet. Available at: <https://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier3&cid=1228772053996>. editor. Accessed: August 21, 2017.
16. CMS. Detailed methods of the 2015 supplemental quality and resource use reports (QRURs). 2016. Available at: <https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/PhysicianFeedbackProgram/Downloads/2015-SQRUR-Detailed-Methods.pdf>. Accessed August 21, 2017.
17. CMS. Hospital IQR program & hospital VBP program: FY 2018 Medicare spending per beneficiary (MSPB) Hospital-specific report (HSR) overview. 2017. Available at: http://www.qualityreportingcenter.com/wp-content/uploads/2017/05/VBPIQR_FY2018_CBMHSRRelease_MSPB_Slides_NPC_05312017_vFINAL5085.31.17-1.pdf. Accessed October 6, 2017.
18. Shih T, Cole AI, Al-Attar PM, et al. Reliability of surgeon-specific reporting of complications after colectomy. *Ann Surg*. 2015;261:920-925.
19. Adams JL, Mehrotra A, Thomas JW, McGlynn EA. Physician cost profiling—reliability and risk of misclassification. *New Engl J Med*. 2010;362:1014-1021.
20. Hayward RA, Heisler M, Adams J, Dudley RA, Hofer TP. Overestimating outcome rates: statistical estimation when reliability is suboptimal. *Health Serv Res*. 2007;42:1718-1738.
21. Markovitz AA, Ellimootil C, Sukul D, et al. Risk adjustment may lessen penalties on hospitals treating complex cardiac patients under medicare's bundled payments. *Health Affairs (Project Hope)*. 2017;36:2165-2174.
22. Mendez CM, Harrington DW, Christenson P, Spellberg B. Impact of hospital variables on case mix index as a marker of disease severity. *Popul Health Manage*. 2014;17:28-34.

EDITORIAL COMMENT



The Centers for Medicare and Medicaid Services (CMS), the nation's largest single payer system, strives to provide quality care to vulnerable populations in a cost-effective manner.¹ Historically, care provision has been a numbers game where providers and health systems were incentivized based on the quantity of care. Value based care programs are evidence of a paradigm shift in evaluating care.

The authors importantly evaluate one such program – the Merit-Based Incentive program relative to physician spending. The Medicare Access and CHIP Reauthorization Act of 2015 (MACRA) established the quality payment program.¹ Under this program physicians are incentivized to provide high quality care to patients. It is estimated that starting in 2019 payments will be linked to quality and value under a Merit-based Incentive Payment System (MIPS), however, the complex manner in which the program is set up allows for payments to be affected nearly 4 years.

The authors reviewed average procedure-based payments utilizing Surveillance Epidemiology End Results (SEER) -Medicare data from urologists for 14,585 patients undergoing prostatectomy, nephrectomy, or cystectomy by 1,895 urologists between 2008 and 2012. They identified wide variation in physician-specific payments for the procedures, and unstable physician spending patterns over time. They conclude that cost-based measures in MIPS will be challenged to change behavior and to reliably distinguish between less efficient versus lower quality care.

The authors' raise valid concerns that variations across physician spending will impact efficiency of care delivery and quality metrics. Considering the challenges surrounding cost evaluations especially involving complicated programs such as the MIPS, it is difficult to accurately evaluate such a program in its infancy. CMS

has yet to finalize risk-adjustment of performance measures, an important measure as it relates to the MIPS program. The eight episode-based measures are currently being field-tested and are not included in 2017 or 2018 MIPS performance years.¹ They will then be shared with stakeholders for review and refinement, for use in later years, further complicating assessment of value-based care.

Identifying challenges and limitations of SEER-Medicare databases is important. Williams et al reported that inherent selection-bias because of unmeasured confounding variables limits interpretation of survival data for prostate cancer patients who undergo radical prostatectomy or radiotherapy.² It will be important to hone in on appropriate risk-adjustment measures. As the Institute of Medicine noted, more comprehensive patient-level data, including improved measures of comorbid conditions, will be needed to ensure high-quality cancer care.^{3,4}

Value based care is important for patients as it can translate to higher quality care, by rewarding providers for coordinated care and positive health outcomes. It may be a while before the system can truly be evaluated, however, findings from this study provide important insight into baseline concerns with the MIPS program. As the authors correctly note, this study provides more questions than answers—are the MIPS measures really valid? Are there better measures out there? Does the MIPS evaluation work for other programs? The study team has done an excellent job. However, more is needed before widespread clinical implementation.

Zaina P. Qureshi, PhD, MPH, MS, Simon Kim, MD, MPH, Charles L. Bennett, MD, PhD, MPP, Department of Health Services Policy and Management, Arnold School of Public Health, University of South Carolina, Columbia, South Carolina; The Southern Network on Adverse Reaction (SONAR) project, the South Carolina Center of Economic Excellence for Medication Safety, College of Pharmacy, University of South Carolina, Columbia, South Carolina; WJB Dorn VA Medical Center, Columbia, South Carolina; University Hospitals Cleveland Medical Center, Urology Institute, Cleveland, Ohio; Case Western Reserve University School of Medicine, Case Comprehensive Cancer Center, Cleveland, Ohio; Yale University, Cancer Outcomes Public Policy and Effectiveness Research (COPPER) Center, New Haven, Connecticut; Department of Urology, University Hospital Cleveland Medical Center, Cleveland, Ohio; The Hollings National Cancer Institute Designated Cancer Center, Medical University of South Carolina, Charleston

References

1. Centers for Medicare and Medicaid Services. www.cms.gov. Accessed 7/2/2018.
2. Williams SB, Huo J, Chamie K, Smaldone MC, Kosarek CD, Fang JE, Ynalvez LA, Kim SP, Hoffman KE, Giordano SH, Chapin BF. Discerning the survival advantage among patients with prostate cancer who undergo radical prostatectomy or radiotherapy: The limitations of cancer registry data. *Cancer*. 2017 May 1;123(9):1617-1624. <https://doi.org/10.1002/cncr.30506>. Epub 2017 Jan 18.
3. Institute of Medicine. *Delivering High-Quality Cancer Care: Charting a New Course for a System in Crisis*; Washington, DC: The National Academies Press; 2013 <https://doi.org/10.17226/18359>.
4. Qureshi ZP, Ganz PA, Bennett CL. Improving the Evidence Base for Delivery of High-Quality Cancer Care. *JAMA Oncol*. 2017 Aug 1;3(8):1029-1031. <https://doi.org/10.1001/jamaoncol.2016.6722>.

<https://doi.org/10.1016/j.urology.2018.04.052>
UROLOGY 123: 119, 2019. © 2018 Elsevier Inc.