

© Health Research and Educational Trust
DOI: 10.1111/1475-6773.12206
RESEARCH ARTICLE

The Early Effects of Medicare's Mandatory Hospital Pay-for-Performance Program

Andrew M. Ryan, James F. Burgess Jr., Michael F. Pesko, William B. Borden, and Justin B. Dimick

Objective. To evaluate the impact of hospital value-based purchasing (HVBP) on clinical quality and patient experience during its initial implementation period (July 2011–March 2012).

Data Sources. Hospital-level clinical quality and patient experience data from Hospital Compare from up to 5 years before and three quarters after HVBP was initiated.

Study Design. Acute care hospitals were exposed to HVBP by mandate while critical access hospitals and hospitals located in Maryland were not exposed. We performed a difference-in-differences analysis, comparing performance on 12 incentivized clinical process and 8 incentivized patient experience measures between hospitals exposed to the program and a matched comparison group of nonexposed hospitals. We also evaluated whether hospitals that were ultimately exposed to HVBP may have anticipated the program by improving quality in advance of its introduction.

Principal Findings. Difference-in-differences estimates indicated that hospitals that were exposed to HVBP did not show greater improvement for either the clinical process or patient experience measures during the program's first implementation period. Estimates from our preferred specification showed that HVBP was associated with a 0.51 percentage point reduction in composite quality for the clinical process measures ($p > .10$, 95 percent CI: $-1.37, 0.34$) and a 0.30 percentage point reduction in composite quality for the patient experience measures ($p > .10$, 95 percent CI: $-0.79, 0.19$). We found some evidence that hospitals improved performance on clinical process measures prior to the start of HVBP, but no evidence of this phenomenon for the patient experience measures.

Conclusions. The timing of the financial incentives in HVBP was not associated with improved quality of care. It is unclear whether improvement for the clinical process measures prior to the start of HVBP was driven by the expectation of the program or was the result of other factors.

Key Words. Pay-for-performance, hospitals, econometrics, health services research

Hospital value-based purchasing (HVBP) is the next step in the evolution of pay-for-performance from an appealing concept to a standard element of the U.S. health care system. The idea of pay-for-performance—that payers of health care should explicitly link provider reimbursement with quality or efficiency outcomes—is compelling. Because patients have a limited ability to observe the quality of care that they receive, providers have lacked the incentive to provide sufficiently high-quality care, resulting in suboptimal quality across the health care system (Arrow 1963; Institute of Medicine 2001). In response, numerous public and private payer initiatives have attempted to incentivize higher quality care through pay-for-performance programs (Rosenthal et al. 2004, 2006; Robinson et al. 2009). The Patient Protection and Affordable Care Act established HVBP, making Medicare payment subject to quality performance for all acute care hospitals in the United States.

The effects of pay-for-performance on provider behavior have been extensively studied (Petersen et al. 2006; Mehrotra et al. 2009; Van Herck et al. 2010; Flodgren et al. 2011), although comparatively little research has focused on hospital-based programs (Mehrotra et al. 2009; Van Herck et al. 2010). Much of the evidence base for hospital pay-for-performance comes from the experience of the Centers for Medicare and Medicaid Services (CMS) and Premier Inc. Hospital Quality Incentive Demonstration, implemented from 2003 to 2009. Initial studies found that hospitals receiving financial incentives in the demonstration improved more on clinical process measures than comparison hospitals (Grossbart 2006; Lindenauer et al. 2007). However, subsequent research suggested that the program did not generate sustained improvements in quality and did not improve patient health outcomes (Ryan 2009; Jha et al. 2012; Ryan, Blustein, and Casalino 2012).

The design of HVBP is different from prior programs in important ways. HVBP gives equal weight to both quality improvement and attainment to determine incentive payments, uses financial penalties in addition to rewards, and incentivizes measures of patient experience in addition to clinical quality (Ryan and Blustein 2012). This early evaluation study uses quality perfor-

Address correspondence to Andrew Ryan, Ph.D., Department of Healthcare Policy and Research, Weill Cornell Medical College, New York, NY 10065; e-mail: amr2015@med.cornell.edu. James F. Burgess, Jr., Ph.D., is with the Veterans Affairs Boston Health Care System, US Department of Veteran Affairs, Boston University School of Public Health, Boston, MA. Michael F. Pesko, Ph.D., and William B. Borden, M.D., M.P.H., are also with the Department of Healthcare Policy and Research, Weill Cornell Medical College, New York, NY. Justin B. Dimick, M.D., M.P.H., Center for Healthcare Outcomes and Policy, University of Michigan, Ann Arbor, MI.

mance data from hospitals that were exposed and not exposed to the program to evaluate the impact of HVBP during its initial period of implementation.

METHODS

Design and Study Population

Under HVBP, acute care hospitals—those paid under Medicare’s Inpatient Prospective Payment System—received payment adjustments beginning in October of 2012 based on their performance on 12 clinical process and 8 patient experience measures from July 1, 2011 through March 31, 2012. HVBP is budget neutral, redistributing hospital payment “withholds” from “losing” to “winning” hospitals. These withholds are equal to 1 percent of hospital payments from diagnosis related groups (DRGs) in the initial implementation period. Incentive payments in HVBP are based on a unique approach that incorporates both quality attainment and quality improvement, incentivizing hospitals for incremental improvements and foregoing the all-or-nothing threshold design of other programs.

Hospitals that are not paid prospectively—including critical access hospitals and hospitals located in Maryland—are not eligible for HVBP. Nonetheless, many of these hospitals have routinely reported data on their quality of care to Hospital Compare, Medicare’s public quality reporting initiative. We used a difference-in-differences study design, to compare changes in quality performance between hospitals that were exposed to HVBP with a set of matched comparison hospitals. For these analyses, we used longitudinal data from Hospital Compare consisting of up to 5 years of data prior to the start of HVBP and three quarters following the start of HVBP.

Performance on the clinical process and patient experience measures during the post-HVBP implementation period was publicly reported on Hospital Compare for Acute Care Hospitals, but not for comparison hospitals. To address this issue, we imputed quality performance for the comparison hospitals during the postintervention period using data from overlapping periods of hospital discharges. For the patient experience measures, Hospital Compare reported the number of achievement points that were received by each hospital participating in HVBP, but it did not report the actual level of performance. We converted hospitals’ achievement points into performance levels using published information on score calculation (Centers for Medicare and Medicaid Services 2011). Appendix A provides a detailed description of these procedures.

Consistent with the rules for HVBP, hospital performance for clinical process measures with denominators of less than 10 were not considered in our analysis. Hospitals in US territories, hospitals reporting data for less than four measures (with denominators of 10 or greater), hospitals with missing data in any period, and acute care hospitals that were otherwise not eligible for participation in the first implementation period were excluded from the analysis. Department of Veterans Affairs hospitals did not report clinical process performance data for a sufficiently long period on Hospital Compare, nor did they report patient experience data, and were therefore excluded from all analyses.

Outcomes

Our two study outcomes are clinical process performance and patient experience performance. Data for these outcomes were downloaded from Hospital Compare (www.hospitalcompare.hhs.gov). To assess clinical process performance, we used data on the 12 measures that were incentivized in the first year of HVBP (Appendix A, Table A1). For these measures, performance scores range from 0 to 100 and can be interpreted as the percentage of opportunities for providing recommended care that were actually provided. To assess patient experience performance, we used data on the 8 measures that were incentivized in the initial performance period of HVBP (Appendix A, Table A1). Following the scoring methodology for HVBP, performance on the patient experience measures was assessed as the percentage of patients reporting “always” to each of the questions (e.g., patients who reported that their doctors “always” communicated well). The one exception is the “Overall rating” measure, in which performance was assessed as the percentage of patients that gave the hospital a rating of 9 or 10 (on a 10-point scale). For both the clinical process and patient experience domains, we created composite measures as the un-weighted mean performance of all reported measures that met denominator requirements.

Matching

Hospitals that were exposed to HVBP tended to be larger, have more admissions, have a higher likelihood of being accredited by the Joint Commission, have a higher likelihood of being a member of the Council of Teaching Hospitals, have higher preintervention performance on the incentivized clinical process measures, and to have lower preintervention

performance on the incentivized patient experience measures (Table 1). Our analysis also found that preintervention trends in clinical process performance were different between hospitals exposed and not exposed to HVBP (Appendix A, Table A2). This violation of the “parallel trends” assumption poses a serious challenge to estimating the impact of policies using difference-in-differences (Angrist and Pischke 2009).

To address this issue, we evaluated the effect of HVBP against a matched cohort of hospitals that were not eligible for HVBP. Recent research suggests that—in the context of difference-in-differences estimation—matching can result in more accurate point estimates and statistical inference when treatment and comparison groups differ on preintervention levels or trends. Statistical matching is related to other methods that have been developed to optimally choose comparison groups to estimate the effect of policies (Abadie, Diamond, and Hainmueller 2010).

We used propensity score matching to match hospitals that were exposed to HVBP to comparison hospitals. We used lagged levels of the outcomes as our only matching variables.¹ Our strategy matched each hospital that was exposed to HVBP to a single nonexposed hospital, employing one-to-one matching with replacement, calipers of .01, and enforcing common support (Smith and Todd 2005). Matching was performed separately for the clinical process and patient experience domains. After matching, our analysis of clinical process performance included 2,801 hospitals that were exposed to HVBP and 240 comparison hospitals, each with six observations. The analysis of patient experience performance included 2,779 hospitals that were exposed to HVBP and 284 comparison hospitals, each with five observations. Figure 1 shows that our matching strategy created a comparison cohort with clinical process performance in the preintervention period that closely mirrored that of hospitals exposed to HVBP. The same was true for the patient experience domain (Appendix A, Figure A6). For both outcomes, trends in preintervention performance were not statistically different after matching (Appendix A, Table A2).

Statistical Analysis

To test the impact of HVBP, we estimated the following equation for hospital j at time t among the propensity score matched sample:

$$\text{Quality}_{jt} = b_0 + b_1 \text{Post}_t + \delta(\text{Post}_t \cdot \text{HVBP}_j) + b_2 u_j + e_{jt} \quad (1)$$

Table 1: Characteristics of Hospital Cohorts

	<i>Hospital Cohort</i>					
	<i>All Included Hospitals</i>		<i>Clinical Process Performance Match</i>		<i>Patient Experience Match</i>	
	<i>Hospitals Exposed to HVBP</i>	<i>Comparison Hospitals</i>	<i>Hospitals Exposed to HVBP</i>	<i>Comparison Hospitals</i>	<i>Hospitals Exposed to HVBP</i>	<i>Comparison Hospitals</i>
Hospitals, <i>n</i>	2,873	399	2,801	240	2,779	284
Hospital type, % ^{*,†,‡}						
Acute care	100.0	0.0	100.0	0.0	100.0	0.0
Critical access	0.0	89.2	0.0	83.9	0.0	58.9
Maryland	0.0	10.8	0.0	16.1	0.0	41.1
Owner, % ^{*,†,‡}						
Government-run	17.3	24.6	17.0	22.9	17.7	16.7
For-profit	18.5	5.8	18.2	5.7	17.5	7.0
Not-for-profit	64.2	69.7	64.7	71.4	64.8	76.3
Region, % ^{*,†,‡}						
Northeast	17.2	9.9	17.5	13.3	17.0	5.8
Southeast	33.2	45.9	33.3	51.9	33.7	61.7
Midwest	8.6	2.4	8.3	0.3	8.7	1.4
South central	13.3	4.5	13.0	3.4	13.1	3.5
North central	8.3	23.7	8.4	16.2	8.2	13.8
Mountain	6.7	7.5	6.7	7.0	6.8	5.9
Pacific	12.7	6.1	12.7	7.9	12.4	8.0
Teaching affiliation, % ^{*,†,‡}	32.3	11.7	32.9	16.1	33.2	21.6
Council of Teaching Hospitals, % ^{*,†,‡}	9.2	2.1	9.4	2.9	9.5	5.2
Joint Commission accredited, % ^{*,†,‡}	90.4	39.7	90.8	33.9	91.5	73.5
Number of beds, mean ^{*,†,‡}	227	69	231	82	238	130
Number of admissions, mean ^{*,†,‡}	10,855	2,844	11,063	3,746	11,252	7,339
Percentage admissions						
Medicare patients, mean ^{*,†,‡}	44.5	54.8	44.5	52.6	45.0	50.0
Medicaid patients, mean ^{*,†,‡}	17.6	13.8	17.6	14.8	18.9	17.0

continued

Table 1: *Continued*

	<i>Hospital Cohort</i>					
	<i>All Included Hospitals</i>		<i>Clinical Process Performance Match</i>		<i>Patient Experience Match</i>	
	<i>Hospitals Exposed to HVBP</i>	<i>Comparison Hospitals</i>	<i>Hospitals Exposed to HVBP</i>	<i>Comparison Hospitals</i>	<i>Hospitals Exposed to HVBP</i>	<i>Comparison Hospitals</i>
Pre-HVBP clinical process performance, mean*	89.4	88.6	89.5	89.0	—	—
Pre-HVBP patient experience performance, mean*	68.6	73.3	—	—	68.6	68.7

Note. Exhibit includes hospitals that are included in either the clinical process or patient experience analysis.

* $p < .05$ for test of difference across all included hospitals.

† $p < .05$ for test of difference across the process performance matched hospitals.

‡ $p < .05$ for test of difference across the patient experienced matched hospitals.

HVBP, hospital value-based purchasing.

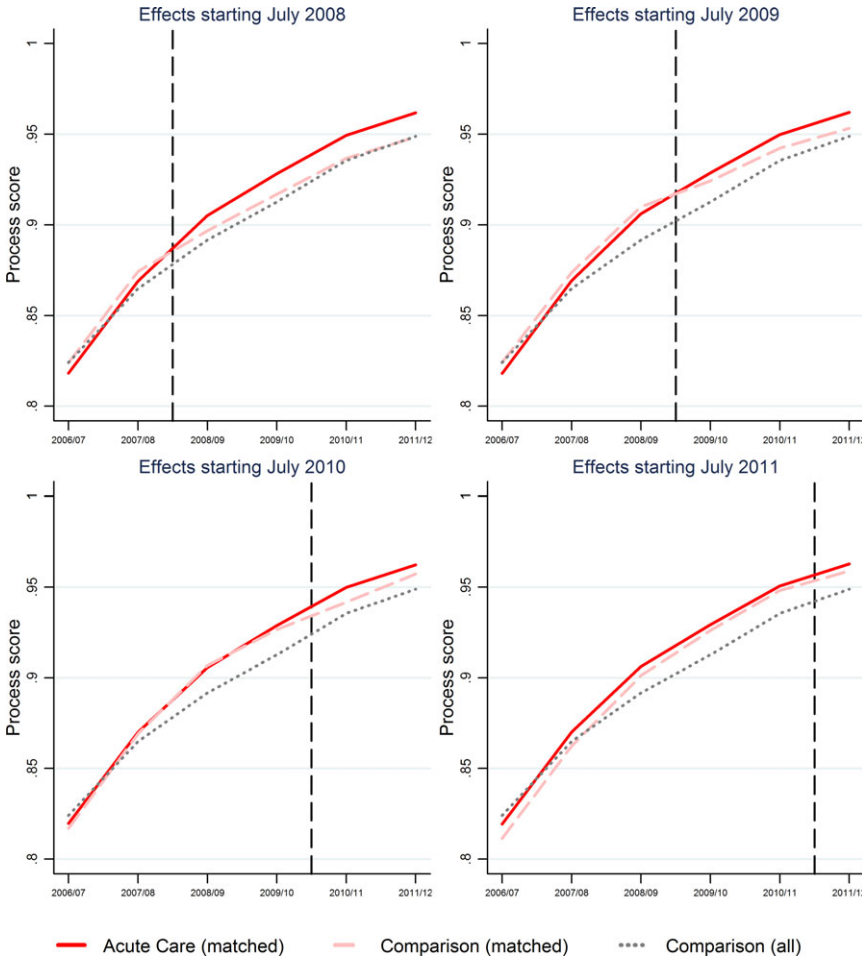
Where *Quality* is composite quality performance for a given domain, *HVBP* is a dummy variable indicating that a hospital was exposed to HVBP, *Post* is dummy variable equal to 1 after the start of HVBP, and *u* is a vector of hospital fixed effects. In equation (1), δ provides the difference-in-differences estimate of the effect of HVBP on quality. The equation was estimated separately for process and patient experience measures.

For process measures, we estimate an additional specification in which we adjusted for variation in the measures reported by each hospital: hospitals that reported a greater share of relatively easy-to-achieve measures would have inflated composite scores, and vice versa. To do this, we created a time-varying variable, which is equal to hospitals’ expected performance on the process composite if they had average performance for each measure that they reported (Ryan and Blustein 2011).

To assess the sensitivity of our results across model specifications, we estimated models that included the entire set of comparison hospitals (not just the matched sample). For clinical process performance, we estimated models with and without the control for the mix of measures reported by hospitals.

We then evaluated whether the effects of HVBP varied across the distribution of quality performance. For instance, it is possible that the effects of

Figure 1: Estimated Effects of Hospital Value-Based Purchasing on Clinical Process Performance Assuming Effects Began between July 2008 and July 2011



Note. Controlling for measure mix and using matched sample, the estimated effect of HVBP assuming that effects started in July 2008: 1.13 (95 percent CI: 0.23, 2.03); July 2009: 0.67 (95 percent CI: -0.13, 1.48); July 2010: 0.14 (95 percent CI: -0.60, 0.87); July 2011: -0.51 (95 percent CI: -1.37, 0.34).

HVBP were greater among hospitals with lower initial performance because these hospitals feared being penalized in the program (Ryan 2013). To do this, we re-estimated our models using quantile regression (Koenker and Hallock

2001), estimating the effects of HVBP at the 5th, 10th, 25th, 50th, 75th, 90th, and 95th percentiles of the distribution of both outcomes.

Hospitals may have anticipated the start of HVBP and begun to improve quality performance prior to the commencement of financial incentives. For instance, the report to Congress outlining the plan to implement HVBP was published in November 2007 (Centers for Medicare and Medicaid Services 2007); the Patient Protection and Affordable Care Act, initiating HVBP, was passed in March of 2010 (2010); and the Final Rule for HVBP, establishing the specific performance measures and incentive structure for the program, was published in May of 2011 (Centers for Medicare and Medicaid Services 2011). To address this issue, we estimated a series of models that assumed that the effects of the program began between 1 and 3 years earlier than the financial incentives were initiated (Angrist and Keueger 1991). For this analysis, we used the previously described matching procedure to create a separate match for each alternative program estimate.

We performed sensitivity analysis by excluding hospitals that had previously participated in the Hospital Quality Incentive Demonstration (Ryan 2009). Using the same matching strategy, we also estimated the effect of HVBP on the individual measures incentivized in HVBP, rather than the composite measures.

Standard errors in all models were robust to clustering at the hospital level. All analysis was performed using Stata 12.0.

RESULTS

Table 2 shows the estimates of the effect of HVBP on clinical process and patient experience. It shows no evidence that the HVBP improved clinical process or patient performance during its first implementation period. For the clinical process measures, our preferred specification—which included the propensity score matched sample and the control for hospitals' reported measure mix—indicates that HVBP was associated with a 0.51 percentage point reduction in composite quality ($p > .10$, 95 percent CI: $-1.37, 0.34$). For patient experience, the estimate from the matched sample indicated that HVBP was associated with a 0.30 percentage point reduction in composite quality ($p > .10$, 95 percent CI: $-0.79, 0.19$).

Table 3 shows the quantile regression estimates of the effects of HVBP across the distribution of outcome performance. It shows no evidence that

Table 2: Estimates of Effects of Hospital Value-Based Purchasing on Clinical Process and Patient Experience Performance

Estimate	Specification Features		Model Results			
	Outcome	Comparison Group	Control for Measure Mix	Number of Hospitals	Number of Observations	Effect Estimate (95% CI)
1	Clinical process	All nonexposed hospitals	No	3,080	18,480	0.50 (-0.18, 1.18)
2	Clinical process	All nonexposed hospitals	Yes	3,080	18,480	0.16 (-0.51, 0.82)
3	Clinical process	Propensity score matched	No	3,041	18,246	-0.14 (-1.00, 0.73)
4	Clinical process	Propensity score matched	Yes	3,041	18,246	-0.51 (-1.37, 0.34)
5	Patient experience	All nonexposed hospitals	No	3,170	12,680	0.16 (-0.11, 0.44)
6	Patient experience	Propensity score matched	No	3,063	12,252	-0.30 (-0.79, 0.19)

Note. No effects were significant at $p < .05$. Incentivized Clinical Process of Care Measures: *Acute Myocardial Infarction*: fibrinolytic therapy; primary percutaneous coronary intervention; *Heart Failure*: discharge instructions; *Pneumonia*: blood cultures performed in the emergency department; initial antibiotic selection; *Surgical Care Improvement*: prophylactic antibiotic received; prophylactic antibiotic selection; prophylactic antibiotics discontinued; cardiac patients with controlled 6AM postoperative serum glucose; venous thromboembolism prophylaxis ordered; appropriate venous thromboembolism prophylaxis 24 hours before and after surgery; appropriate continuation of beta blocker. Incentivized patient experience measures: communication with nurses; communication with doctors; responsiveness of hospital staff; pain management; communication about medicines; cleanliness and quietness of hospital environment; discharge information; overall rating of hospital.

Table 3: Quantile Regression Estimates of the Effect of Hospital Value-Based Purchasing on Clinical Process and Patient Experience Performance

Percentile	Effect Estimate (95% CI)	
	Clinical Process	Patient Experience
5th	-0.26 (-1.23, 0.71)	-0.23 (-0.92, 0.46)
10th	-0.45 (-1.03, 0.13)	-0.85 (-2.13, 0.42)
25th	0.11 (-0.76, 0.97)	-0.61 (-1.14, -0.08)
50th	-0.50 (-1.31, 0.32)	-0.09 (-0.71, 0.54)
75th	-0.13 (-0.78, 0.53)	-0.15 (-0.84, 0.55)
90th	-0.24 (-1.27, 0.79)	-0.51 (-1.44, 0.43)
95th	0.35 (-1.76, 2.46)	-0.06 (-1.15, 1.02)

Note. No effects were significant at $p < .05$. Clinical process models include 21,287 observations from 3,041 hospitals; patient experience models include 15,315 observations from 3,063 hospitals. 95% CIs based are based on block-bootstrap standard errors. Estimates are based on models with controls for measure mix and using the matched comparison group.

initially higher or lower performing hospitals improved more in response to the program during the first performance period.

Figure 1 shows the results of our analysis testing whether hospitals that were ultimately exposed to HVBP began to improve quality clinical process performance in advance of program implementation. Effects are shown for the start of financial incentives in HVBP (July 2011), when we assume that HVBP began to affect hospitals 1 year before financial incentives began (July 2010), 2 years before financial incentives began (July 2009), and 3 years before financial incentives began (July 2008). Figure 1 shows some evidence that hospitals did in fact improve clinical process quality in advance of HVBP. In models using the matched sample and controlling for measure mix, the estimated effect of HVBP is +1.13 percentage points ($p < .05$, 95 percent CI: 0.23, 2.03) when assuming that its effects began in July 2008, +0.67 ($p < .10$, 95 percent CI: -0.13, 1.48) when assuming that its effects began in July 2009, and +0.14 ($p > .10$, 95 percent CI: -0.60, 0.87) when assuming that its effects began in July 2010. However, we found no evidence that hospitals improved patient experience performance in advance of HVBP (Appendix A, Table A6).

Analysis of the effect of HVBP on the individual incentivized measures found that the program was significantly associated with improved performance for the two clinical process measures related to pneumonia (*blood cultures performed in the emergency department prior to initial antibiotic received in hospital* and *patients received appropriate initial antibiotic*) (Appendix A, Table A3). However, these effects were driven primarily by differences in

performance between hospitals exposed and not exposed to HVBP prior to the start of the program (Appendix A, Figure A5). HVBP was not associated with improved performance for any of the patient experience measures. Sensitivity analysis that excluded the hospitals that participated in the Hospital Quality Incentive Demonstration yielded nearly identical results (Appendix A, Table A4).

DISCUSSION

This is the first study to evaluate the effect of HVBP on quality in its initial implementation period. We found no evidence that improvement in clinical process or patient experience performance was greater for hospitals exposed to HVBP compared to a matched comparison group of hospitals that were not exposed to HVBP. We also found no evidence that the effect of HVBP varied based on hospitals' initial clinical process or patient experience performance. We did, however, find some evidence that hospitals that were ultimately exposed to HVBP had greater improvement on clinical process performance when we assumed that the effects of HVBP began 3 years prior to the start of financial incentives. Whether this improved performance was driven by the expectation of HVBP, or whether it resulted from other factors, is unclear.

Results from this study are consistent with other evidence from the United States that hospital pay-for-performance programs have resulted in little to no improvement in quality of care (Glickman et al. 2007; Ryan 2009; Ryan and Blustein 2011; Jha et al. 2012; Ryan, Blustein, and Casalino 2012). What this means for the future of HVBP is uncertain. The design of HVBP will evolve in the coming years to increase the magnitude of financial incentives and focus on measures of outcome quality (Ryan and Blustein 2012). However, the number of incentivized measures will also increase over time, potentially diluting the effects of increasing the magnitude of incentives.

Limitations

The comparison hospitals that were not exposed to HVBP are different from acute care hospitals across a number of distinct dimensions. The expectations for quality improvement among the comparison hospitals may therefore be different from acute care hospitals. However, after matching, comparison hos-

pitals and hospitals that were exposed to HVBP had nearly identical levels and trends in quality performance before the start of HVBP. This makes post-HVBP performance for the matched comparison hospitals a good counterfactual for hospitals that were exposed to the program. Maryland hospitals—which were included in the comparison group—were subject to pay-for-performance incentives during the study period that were similar to HVBP (Calikoglu, Murray, and Feeney 2012). However, the timing of pay-for-performance in Maryland (beginning in 2008) was different than that of HVBP, allowing us to identify whether the introduction of HVBP led to incremental improvements for exposed hospitals.

In addition, our study may have been underpowered to detect the effects of HVBP. A power analysis conducted prior to the study found that we had moderate to excellent power to detect program effects for both study outcomes for program effects of 1.0 percentage points and above. Evaluations of other pay-for-performance programs found program effects in this range (Lindenauer et al. 2007).

We also found that performance on the incentivized measures improved for both exposed and nonexposed hospitals, and it is possible that the incentives of HVBP “spilled over” to hospitals that were not exposed to the program, contaminating the comparison hospitals. If true, HVBP may have been more successful at improving quality of care than our study suggests. It is also possible that ceiling effects—the decreased expectation of quality improvement once hospitals approach maximum performance scores—limited our ability to determine the effectiveness of the program, particularly for the clinical process measures. However, evidence of the effectiveness of the program did not vary across levels of initial performance, and sensitivity analysis did not find evidence that HVBP was more likely to improve quality for individual measures with lower initial performance.

Finally, our study assessed quality performance in a three-quarter period following the implementation of HVBP. It may take hospitals longer to respond to the financial incentives of the program. HVBP will continue to evolve over the next several years, and results from this study may not hold across future variations in the program design. For instance, hospitals may be more strongly encouraged to improve quality as the revenue at stake in HVBP increases from 1 percent of Medicare revenue to 2 percent of revenue over the next several years. Research on the long-term impact of HVBP will be critical to policy makers. Nevertheless, as seen by the creation of the Rapid Cycle Evaluation Group by the Center for Medicare and Medicaid Innovation, rapid cycle evaluations to inform future policy are crucially important to

establish expectations for similar programs, and modify the design of these programs accordingly (Shrank 2013).

Considerations for the Future of Hospital Value-Based Purchasing

The design of HVBP was substantially different from prior programs in a number of ways, which raised expectations that it might be more effective in improving quality of care (Ryan and Blustein 2012). On net, however, these design differences did not appear to motivate quality improvement in HVBP's first implementation period. It is possible the magnitude of the financial incentives remained too low to motivate quality improvement (Ryan 2013), particularly given hospitals' competing priorities from Medicare's new Hospital Readmission Reduction Program and other policy reforms. In addition, the complicated nature of the incentive design may have failed to give participating hospitals clear targets for performance, attenuating improvement (Jha 2013).

HVBP is statutorily defined in the Patient Protection and Affordable Care Act and will continue indefinitely. However, CMS has some flexibility to modify its design through the rulemaking process. CMS should continue to experiment with the performance measures and the incentive structure in HVBP until the program is shown to improve quality of care. This includes increasing the financial incentives in the program and identifying performance measures with sufficient room for quality improvement. As HVBP evolves, research should continue to better understand the conditions under which value-based purchasing programs can be effective.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: The authors acknowledge Jayme Mendelsohn for research assistance. This research was funded by the Robert Wood Johnson Foundation under the Changes in Health Care Financing and Organization program (grant 70782). Funding for Dr. Ryan was provided by a career development award from the Agency for Healthcare Research and Quality (K01 HS018546). William Borden does work with the US Department of Health and Human Services. His work on this study was conducted through Weill Cornell Medical College and is not related to the Department of Health and Human Services.

Disclosures: Dr. Dimick was supported by a grant from the National Institute on Aging (R01AG039434). Dr. Dimick is a paid consultant and equity

owner in ArborMetrix, Inc., a company that provides software and analytic services for assessing hospital quality and efficiency.

Disclaimers: None.

NOTE

1. For the clinical process measures, matching used the 1st through 5th lags of composite quality. For the patient experience measures, matching used the 2nd through 4th lags of composite quality. While the results were not sensitive to that were specified for matching, matching on the 2nd through 4th lags resulted in a better fit.

REFERENCES

- Abadie, A., A. Diamond, and J. Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association* 105 (490): 493–505.
- Angrist, J. D., and A. B. Keueger. 1991. "Does Compulsory School Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics* 106 (4): 979–1014.
- Angrist, J. D., and J. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.
- Arrow, K. J. 1963. "Uncertainty and the Welfare Economics of Medical-Care." *American Economic Review* 53 (5): 941–73.
- Calikoglu, S., R. Murray, and D. Feeney. 2012. "Hospital Pay-For-Performance Programs in Maryland Produced Strong Results, Including Reduced Hospital-Acquired Conditions." *Health Affairs (Millwood)* 31 (12): 2649–58.
- Centers for Medicare and Medicaid Services. 2007. *Report to Congress: Plan to Implement a Medicare Hospital Value-Based Purchasing Program*. Washington, DC: U.S. Department of Health and Human Services.
- Centers for Medicare and Medicaid Services. 2011. "Medicare Program; Hospital Inpatient Value-Based Purchasing Program. Final Rule." *Federal Register* 76 (88): 26490–547.
- Flodgren, G., M. P. Eccles, S. Shepperd, A. Scott, E. Parmelli, and F. R. Beyer. 2011. "An Overview of Reviews Evaluating the Effectiveness of Financial Incentives in Changing Healthcare Professional Behaviours and Patient Outcomes." *Cochrane Database of Systematic Reviews* 7: CD009255.
- Glickman, S. W., F. S. Ou, E. R. DeLong, M. T. Roe, B. L. Lytle, J. Mulgund, J. S. Rumsfeld, W. B. Gibler, E. M. Ohman, K. A. Schulman, and E. D. Peterson. 2007. "Pay for Performance, Quality of Care, and Outcomes in Acute Myocardial Infarction." *Journal of the American Medical Association* 297 (21): 2373–80.

- Grossbart, S. R. 2006. "What's the Return? Assessing the Effect of "Pay-For-Performance" Initiatives on the Quality of Care Delivery." *Medical Care Research and Review* 63 (1 Suppl): 29S–48S.
- Institute of Medicine. 2001. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: The National Academies Press.
- Jha, A. K. 2013. "Time to Get Serious about Pay for Performance." *Journal of the American Medical Association* 309 (4): 347–8.
- Jha, A. K., K. E. Joynt, E. J. Orav, and A. M. Epstein. 2012. "The Long-Term Effect of Premier Pay for Performance on Patient Outcomes." *New England Journal of Medicine* 366 (17): 1606–15.
- Koenker, R., and K. F. Hallock. 2001. "Quantile Regression." *Journal of Economic Perspectives* 15 (4): 143–56.
- Lindenaier, P. K., D. Remus, S. Roman, M. B. Rothberg, E. M. Benjamin, A. Ma, and D. W. Bratzler. 2007. "Public Reporting and Pay for Performance in Hospital Quality Improvement." *New England Journal of Medicine* 356 (5): 486–96.
- Mehrotra, A., C. L. Damberg, M. E. Sorbero, and S. S. Teleki. 2009. "Pay for Performance in the Hospital Setting: What Is the State of the Evidence?" *American Journal of Medical Quality* 24 (1): 19–28.
- Patient Protection and Affordable Care Act. 2010. H.R 3590. Public Law 111–48.
- Petersen, L. A., L. D. Woodard, T. Urech, C. Daw, and S. Sookanan. 2006. "Does Pay-for-Performance Improve the Quality of Health Care?" *Annals of Internal Medicine* 145 (4): 265–72.
- Robinson, J. C., S. M. Shortell, D. R. Rittenhouse, S. Fernandes-Taylor, R. R. Gillies, and L. P. Casalino. 2009. "Quality-Based Payment for Medical Groups and Individual Physicians." *Inquiry* 46 (2): 172–81.
- Rosenthal, M. B., R. Fernandopulle, H. R. Song, and B. Landon. 2004. "Paying for Quality: Providers' Incentives for Quality Improvement." *Health Affairs (Millwood)* 23 (2): 127–41.
- Rosenthal, M. B., B. E. Landon, S. L. Normand, R. G. Frank, and A. M. Epstein. 2006. "Pay for Performance in Commercial HMOs." *New England Journal of Medicine* 355 (18): 1895–902.
- Ryan, A. M. 2009. "Effects of the Premier Hospital Quality Incentive Demonstration on Medicare Patient Mortality and Cost." *Health Services Research* 44 (3): 821–42.
- . 2013. "Will Value-Based Purchasing Increase Disparities in Care?" *New England Journal of Medicine* 369 (26): 2472–4.
- Ryan, A. M., and J. Blustein. 2011. "The Effect of the MassHealth Hospital Pay-for-Performance Program on Quality." *Health Services Research* 46 (3): 712–28.
- Ryan, A., and J. Blustein. 2012. "Making the Best of Hospital Pay for Performance." *New England Journal of Medicine* 366 (17): 1557–9.
- Ryan, A. M., J. Blustein, and L. P. Casalino. 2012. "Medicare's Flagship Test of Pay-For-Performance Did Not Spur More Rapid Quality Improvement among Low-Performing Hospitals." *Health Affairs (Millwood)* 31 (4): 797–805.

- Shrank, W. 2013. "The Center For Medicare And Medicaid Innovation's Blueprint for Rapid-Cycle Evaluation of New Care and Payment Models." *Health Affairs (Millwood)* 32 (4): 807–12.
- Smith, J. A., and P. E. Todd. 2005. "Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators?" *Journal of Econometrics* 125: 305–53.
- Van Herck, P., D. De Smedt, L. Annemans, R. Remmen, M. B. Rosenthal, and W. Sermeus. 2010. "Systematic Review: Effects, Design Choices, and Context of Pay-For-Performance in Health Care." *BMC Health Services Research* 10: 247.

SUPPORTING INFORMATION

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

Appendix SA1: Author Matrix.

Data S1: Methodological Details.