International Journal for Quality in Health Care, 2017, 29(2), 222–227
doi: 10.1093/intqhc/mzw159
Advance Access Publication Date: 12 April 2017
Article

Article

Pay-for-performance reduces healthcare spending and improves quality of care: Analysis of target and non-target obstetrics and gynecology surgeries

SEUNG JU KIM[1,2], KYU-TAE HAN[1,2], SUN JUNG KIM[3], and EUN-CHEOL PARK[2,4]

[1]Department of Public Health, Graduate School, Yonsei University, Seoul, Republic of Korea, [2]Institute of Health Services Research, Yonsei University, Seoul, Republic of Korea, [3]Department of Health Administration, Soonchunhyang University, Chungnam, Republic of Korea, and [4]Department of Preventive Medicine, Yonsei University College of Medicine, Seoul, Republic of Korea

Address reprint requests to: Eun-Cheol Park, Department of Preventive Medicine & Institute of Health Services Research, Yonsei University College of Medicine, 50 Yonsei-ro, Seodaemun-gu, Seoul 120-752, Republic of Korea. Tel: +82-2-2228-1862; Fax: +82-02-392-8133; E-mail: ecpark@yuhs.ac

Editorial Decision 18 December 2016; Accepted 11 April 2017

Abstract

Objective: In Korea, the Value Incentive Program (VIP) was first applied to selected clinical conditions in 2007 to evaluate the performance of medical institutes. We examined whether the condition-specific performance of the VIP resulted in measurable improvement in quality of care and in reduced medical costs.
Design: Population-based retrospective observational study.
Setting: We used two data set including the results of quality assessment and hospitalization data from National Health Claim data from 2011 to 2014.
Participants: Participants who were admitted to the hospital for obstetrics and gynecology were included. A total of 535 289 hospitalizations were included in our analysis.
Methods: We used a generalized estimating equation (GEE) model to identify associations between the quality assessment and length of stay (LOS). A GEE model based on a gamma distribution was used to evaluate medical cost. The Poisson regression analysis was used to evaluate readmission.
Main Outcome Measures: The outcome variables included LOS, medical costs and readmission within 30 days.
Results: Higher condition-specific performance by VIP participants was associated with shorter LOSs, decreases in medical cost, and lower within 30-day readmission rates for target and non-target surgeries. LOS and readmission within 30 days were different by change in quality assessment at each medical institute.
Conclusions: Our findings contribute to the body of evidence used by policy-makers for expansion and development of the VIP. The study revealed the positive effects of quality assessment on quality of care. To reduce the between-institute quality gap, alternative strategies are needed for medical institutes that had low performance.

Key words: quality improvement, quality indicators, patient outcomes, health policy, readmission

Introduction

The growth of health expenditures and improving quality of care is of major concern to policy-makers, healthcare providers and other stakeholders. One method used to improve quality of care is pay-for-performance (P4P), which measures quality indicators [1–3]. Many countries have adopted the P4P program and provide financial incentives or penalties to medical institutes according to their efforts to improve healthcare quality [4, 5].

In Korea, healthcare system was operated by single payer under the National Health Insurance Service (NHIS), and reimbursed through a fee-for-service system to all disease except to seven disease groups which were applied into Diagnosis-Related Groups (DRGs)-based payment system. In addition, quality assessment for healthcare providers was introduced with the reformation of the NHI Act in 2000. Under the Act, the Health Insurance Review and Assessment Service (HIRA) was formed to evaluate the performance of medical institutions [6, 7]. After several years of evaluation of the performance of medical institutes, the Value Incentive Program (VIP) was introduced in June 2007 in the areas of acute myocardial infarction (AMI) and cesarean section (CS) in only tertiary hospitals; because these hospitals already satisfied certain standards for healthcare quality assessment. During the pilot program for tertiary hopitals, the most remarkable improvement is shown on the process of treatment in AMI; percentage of patients who were given fibrinolytic therapy (within 60 min of hospital arrival) and percutaneous coronary intervention (within 120 min of hospital arrival) had been increased as 20.9% and 10.7%, respectively. While a slight improvement was observed with respect to 30 days mortality (2007: 7.9%, 2009: 7.7%). The rate of CS was also decreased from 34.3% in 2007 to 33.6% in 2009. Based on an evaluation of a VIP pilot system, hospitals that had high performance received incentives or penalties beginning in 2010. In 2012, the VIP was officially expanded, including prophylactic use of antibiotics in 11 surgeries, to other medical institutes including hospitals, general hospitals and tertiary hospitals. Each year, the HIRA has increased the number of items used to measure performance of medical institutes in disease, surgery and medicine to improve quality of care and reduce quality gaps among healthcare institutes, and to provide information about medical institutions to the public. Incentives or penalties were given to medical institutes with high or low quality grades, respectively, based on 3 years of performance [6]. The first year, the quality grade results were available to the public, but only included the highest two grades received in the quality assessment. The subsequent year, the quality grade results for all medical institutes were available to the public and incentives were awarded to medical institutes that had received a higher quality assessment grade or had maintained the highest grade. The final year (Year 3), penalties were given to the medical institutes with grades in lower thresholds for each quality assessment. Based on the 2012 results, in 2013 the VIP was applied to medical institutes that had upgraded to a higher grade or maintained high grades for selected conditions (AMI, CS and prophylactic use of antibiotics). In 2014, AMI and CS were excluded from the VIP, and the treatment of acute stroke, prophylactic use of antibiotics and the use of medications (antibiotic prescription rate, prescription rate of injections, numbers of medications) were added to the VIP. Improvements in quality of care occurred in selected disease areas after adoption of the VIP in Korea [7, 8]. However, it is unclear whether other quality measures also result in actual improvement in quality of care during care of patients. Criteria for measuring the use of surgical antibiotic prophylaxis have not been included in measurement of patient outcomes.

Previous research results reveal that incentive programs are associated with improvements in quality of care [9–11]. However, some results also suggest that these programs do not result in better quality of care [12–14]. Numerous researchers have looked at indicators for measuring quality [15–17]. In addition, researches have suggested that the prophylactic use of antibiotics is associated with lower complication rates [18, 19]. Pioneering studies on the effects of healthcare providers' attitudes towards the VIP have been performed in Korea [20, 21]. Other studies have examined the effects of the quality of prophylactic use of antibiotics on patient outcome in smaller populations [22]. Less is known about the effects of prophylactic use of antibiotics on quality of care.

We evaluated two aspects of the quality assessment of prophylactic use of antibiotics. First, we examined whether it resulted in improvement in quality of care, including length of stay (LOS), medical cost and readmission within 30 days, and whether similar results were observed for other surgeries that had not been selected for the targeted condition. We also evaluated the association between quality assessment and quality of care by examining changes in organizational performance.

Methods

Data collection

Quality assessment of the prophylactic use of antibiotics was applied to 15 surgery types in 2014. The selected surgeries were applied to different reimbursement systems. To minimize potential effects of reimbursement system on quality of care [23, 24], we selected obstetrics and gynecology surgery (OBGY) for the study population because most of the OBGY surgeries were applied to the DRGs-based payment system. We used NHI claim data (2011–14) to investigate the effects of the quality assessment on quality of care. The criteria for excluding patients from the study population consisted of characteristics that excluded patients from the DRGs-based payment system and quality assessment. We excluded medicaid patients who were not suitable for the DRGs-based payment system. We also excluded both patients <18 years of age and CS patients with multiple fetuses, which were not included in quality assessment. In addition, quality is evaluated in hospitals, general hospitals and tertiary hospitals, we excluded unsuitable hospital such as clinics. We included surgeries for hysterectomy, CS, uterine procedures and adnexa procedures; surgery selected for quality assessment was classified as target surgery and non-selection was classified as non-target surgery. A total of 535 289 hospitalizations at 506 hospitals were included in our analysis.

Variables

Quality assessment of the prophylactic use of antibiotics was evaluated using six indicators every 2 years: (i) the rate of prophylactic antibiotic use before 1 h to skin incision, (ii) the rate of aminoglycoside antibiotic use, (iii) the rate of third-generation cephalosporin antibiotic use, (iv) the rate of use of multiple antibiotics, (v) prescription rate of antibiotics at discharge, and (vi) total duration of prescription of antibiotics. A composite quality score of these six indicators was calculated for each hospital. Quality score grades ranged from 0 to 5 (0: exclusion of grade; 5: high score medical institutes; 1: lower score medical institute). We considered the quality score grade to be a continuous variable because it was based on the total quality score. We defined the condition-specific performance as quality assessment for prophylactic use of antibiotics;

because we selected specific condition in VIP program. Furthermore, we added an additional item of quality measurement of antibiotic use to adjust for changes in quality measurement in each medical institute. We assigned the change in quality measurement of antibiotic use to one of three groups: upgrade in grade, decline in grade and no change in grade.

The outcome variables included LOS, medical cost and readmission within 30 days. LOS was measured using each patient's dates of admission and discharge. We used a log transformation for LOS to reflect the original scale of the skewed data and to measure changes in the dependent variable in response to percentage changes in the explanatory variable [25–27]. Medical cost was evaluated by calculating the total cost (patient and insurer cost) of the entire treatment period (hospitalization and outpatient), including the 30-day period after discharge from the hospital within 30 days. To account for variations in inflation, we adjusted the price forward to 2011 KRW (Korean Won) using an annual conversion factor. We defined readmission within 30 days as a patient who was readmitted with the same diagnosis as the previous admission or with one or more complications (complications of surgical and medical care/ ICD 10: T80–T88).

Hospital characteristics, such as hospital type (hospital, general hospital, tertiary hospital), teaching status (teaching, non-teaching), hospital location (urban, rural), number of beds and human resources (numbers of doctors and nurses per 100 beds) were included in the analysis. To minimize the confounding effects of differences across hospitals, we adjusted for the volume per hospital and for the case mix index (CMI). The patient characteristics included in the analysis were patient ID, age, Charlson comorbidity index (CCI), use of laparoscopy (Yes, No) and year.

Statistical analysis

The distribution of each categorical variable was examined using an analysis of frequencies and percentages; χ^2 tests were performed to examine associations with categorical variables. Analysis of variance was also performed to compare mean values and standard deviations for continuous variables. A generalized estimating equation (GEE) model was used to evaluate the effects of quality assessment on LOS, medical cost and readmission. The correlation structure was modeled as an exchangeable correlation to determine the repeated outcome variables [28, 29]. We used a gamma generalized linear model based on the log link function to evaluate medical cost [27]. Because of the low readmission prevalence, a GEE with a Poisson regression analysis was used to examine associations between readmission and quality assessment [30]. Subgroup analyses were performed based on change in quality assessment of antibiotic use. All statistical analyses were performed using SAS version 9.4 (SAS Institute, Inc., Cary, NC, USA). A P-value < 0.05 indicated that a result was statistically significant.

Results

The data used in this study consisted of 535 289 cases at 506 hospitals. Among the OBGY surgeries, mean LOS was longest for hysterectomy (6.97; SD: ± 2.68) and shortest for uterine procedure (5.24; SD: ± 2.35). Total medical cost was lowest for CS (1 506 605; SD: ± 427 230 KRW) and highest for hysterectomy (2260,113; SD: ± 622 666 KRW). Frequency of readmission was highest for hysterectomy (n = 2920, 2.8%) and lowest for CS (n = 797, 0.3%) (Table 1).

The results of the regression analysis using a GEE revealed that LOS was significantly decreased with higher condition-specific performance for both target and non-target surgeries. Higher condition-specific performance was significantly associated with reduction in medical cost. An increase in quality grade was associated with a lower risk of readmission within 30 days for CS and for uterine procedure (Table 2).

The result for the subgroup analysis by change in quality assessment of antibiotic use indicated the presence of a general trend of a decrease in LOS at medical institutes that did not change quality grade. Higher condition-specific performance was associated with a reduction in medical cost regardless of whether there was a change in quality assessment. Readmission differed by the change in quality assessment of antibiotic use; patients had a lower risk of readmission when they were treated at medical institutes that had not changed quality grade (Table 3).

Discussion

In this study, we evaluated the associations between quality assessment and quality of care, and whether there were actual improvements in patient care. The study revealed that higher condition-specific performance was associated with shorter LOSs and decreases in medical costs and readmission within 30 days for both target and non-target surgeries. Our findings were similar to the findings of previous studies; higher performance in the quality assessment was associated with improvement in quality of care [31, 32]. Other studies have revealed that higher performance is associated with lower values for prevalence of postoperative infection and complications [18, 19]. Medical institutes with a higher quality assessment grade result had better management of antibiotics during preoperative and postoperative patient care. Quality assessment criteria include the selection, timing and period in use of antibiotics, which have been recommended as indices associated with decreases in the prevalence of postoperative infections. As a result of better antibiotic management, periods of patient care during hospitalization were shortened and decreases in readmission within 30 days resulted due to the lower complication rates. Suitable care during preoperative and postoperative periods also decreases total healthcare costs per patient. In our study, the size of effect was not large in terms of the decreases in medical costs per surgery, but the results indicated that a higher quality assessment grade was associated with a reduction in health expenditures. Similar results were observed for the target and non-target surgeries. One explanation for these results is that the information was available to the public. Because the public report of the quality assessment results was available and presented an index of quality for each medical institute, there would also be an effect on the use of antibiotics for other non-target surgeries. However, these results might be unrelated to the specific variations in quality assessment that occurred at each medical institute. Further evaluation by change in quality assessment is needed.

The subgroup analysis revealed that the result for medical cost was similar to the result that higher condition-specific performance was associated with decreases in healthcare cost. However, the results were different by change in quality assessment for LOS and readmission within 30 days after discharge. There was a general decreased trend for LOS at medical institutes with no change in quality assessment grade. For medical institutes that upgraded their quality grade, the LOS was decreased except CS, with a higher condition-specific performance. Although it was not statistically significant, there was an increase in LOS for hysterectomy and uterine procedure at medical institutes with decreases in the quality assessment grade. Higher condition-specific performance was associated with reductions in readmission at medical institutes with no changes

Table 1 General characteristics of participants and hospitals (unit: N/M, %, SD)

	Hysterectomy	CS	Uterine procedure	Adnexa procedure	*P*-value
Outcome variables					
LOS	6.97 ± 2.68	6.74 ± 2.46	5.24 ± 2.35	5.67 ± 2.46	<0.0001
Medical cost	2 260 113 ± 622 666	1 506 605 ± 427 230	1 883 545 ± 667 023	1 996 480 ± 532 235	<0.0001
Readmission within 30 days					
Yes	2920 (2.8)	797 (0.3)	839 (1.3)	1186 (1.4)	<0.0001
No	100 351 (97.2)	279 912 (99.7)	63 001 (98.7)	86 283 (98.6)	
Hospital characteristics (*n* = 506)					
Quality assessment of prophylactic antibiotic use	4.34 ± 1.20	3.71 ± 1.45	4.11 ± 1.40	4.31 ± 1.27	<0.0001
Change in quality assessment of antibiotic use					
Upgrade in grade (*n* = 69)	12 623 (12.2)	37 079 (13.2)	6438 (10.1)	10 304 (11.8)	<0.0001
Decline in grade (*n* = 115)	9232 (8.9)	30 138 (10.7)	5129 (8.0)	7320 (8.4)	
Not changed (*n* = 322)	81 416 (78.8)	213 492 (76.1)	52 273 (81.9)	69 845 (79.9)	
Type of hospital					
Hospital (*n* = 231)	23 420 (22.7)	201 059 (71.6)	22 990 (36.0)	21 522 (24.6)	<0.0001
General hospital (*n* = 231)	43 289 (41.9)	46 778 (16.7)	23 034 (36.1)	33 965 (38.8)	
Tertiary hospital (*n* = 44)	36 562 (35.4)	32 872 (11.7)	17 816 (27.9)	31 982 (36.6)	
Teaching status					
Teaching (*n* = 144)	72 114 (69.8)	70 019 (24.9)	36 970 (57.9)	59 818 (68.4)	<0.0001
Non-teaching (*n* = 362)	31 157 (30.2)	210 690 (75.1)	26 870 (42.1)	27 651 (31.6)	
Hospital location					
Urban (*n* = 466)	101 376 (98.2)	278 262 (99.1)	63 213 (99.0)	85 938 (98.3)	<0.0001
Rural (*n* = 40)	1895 (1.8)	2447 (0.9)	627 (1.0)	1531 (1.8)	
Number of beds	813.70 ± 667.82	353.31 ± 525.56	657.47 ± 666.09	815.02 ± 708.16	<0.0001
Number of doctors per 100 beds	29.00 ± 15.62	19.20 ± 12.36	27.11 ± 15.29	29.59 ± 16.20	<0.0001
Number of nurses per 100 beds	51.16 ± 25.11	35.98 ± 28.64	50.67 ± 27.39	52.21 ± 26.22	<0.0001
Volume per hospital	6956.57 ± 5565.86	4543.50 ± 4380.49	6810.01 ± 5808.49	7153.82 ± 5791.06	<0.0001
CMI	1.39 ± 0.16	1.38 ± 0.11	1.44 ± 0.18	1.39 ± 0.17	<0.0001
Patient characteristics					
Age	48.85 ± 8.71	32.20 ± 4.11	40.43 ± 7.45	37.18 ± 12.02	<0.0001
CCI	1.42 ± 1.01	0.04 ± 0.21	0.65 ± 0.75	0.63 ± 1.04	<0.0001
Use of laparoscopy					
Yes	65 008 (63.0)	(0.0)	41 005 (64.2)	76 535 (87.5)	<0.0001
No	38 263 (37.1)	280 709 (100.0)	22 835 (35.8)	10 934 (12.5)	
Year					
2011.07–2012.06	36 941 (35.8)	93 814 (33.4)	19 966 (31.3)	29 747 (34.0)	<0.0001
2012.07–2013.06	34 763 (33.7)	94 250 (33.6)	20 681 (32.4)	29 014 (33.2)	
2013.07–2014.07	31 567 (30.6)	92 645 (33.0)	23 193 (36.3)	28 708 (32.8)	
Total	103 271 (19.3)	280 709 (52.4)	63 840 (11.9)	87 469 (16.3)	

Cost: Adjusted for gross price inflation—that is, as if the gross-to-cost ratio had stayed constant since 2011, and included hospitalization and outpatient care. CCI, Charlson comorbidity index.

Table 2 Associations between condition-specific performance and LOS, medical cost and readmission within 30 days

		LOS			Medical cost				Readmission within 30 days		
		Estimate[a]	Standard error	*P*-value	Estimate	RR	95% CI		RR	95% CI	
Target surgery	Hysterectomy	−0.0191	0.0010	<0.0001	−0.0085	0.992	0.990	0.993	0.967	0.925	1.011
	CS	−0.0030	0.0003	<0.0001	−0.0029	0.997	0.997	0.998	0.913	0.846	0.986
Non-target surgery	Uterine procedure	−0.0052	0.0016	0.0013	−0.0047	0.995	0.994	0.997	0.918	0.849	0.993
	Adnexa procedure	−0.0144	0.0011	<0.0001	−0.0027	0.997	0.996	0.999	1.024	0.958	1.094

Adjusted for change in quality assessment of antibiotic use, type of hospital, teaching status, hospital location, number of beds, number of doctors per 100 beds, number of nurses per 100 beds, volume per hospital, CMI, age, CCI, use of laparoscopy and year, 95% CI: confidence interval.

RR (rate ratio): RR is indicated the results of exponentiated estimates and interpretable as percentage changes.

[a]Estimates are the results of log transformation and interpretable as percentage changes.

in quality grade. However, an increased risk of readmission was observed at medical institutes with a decrease in quality grade. These results suggested that the financial incentives of VIP programs had a positive effect on the quality of care. Rewarding only those high-quality medical institutions that maintain or improve their quality grades could motivate healthcare providers to achieve

Table 3 Subgroup analysis of condition-specific performance on LOS, medical cost and readmission within 30 days, by change in quality assessment of antibiotic use

	LOS			Medical cost				Readmission within 30 days		
	Estimate[a]	Standard error	*P*-value	Estimate	RR	95% CI		OR	95% CI	
Upgrade in grade										
Hysterectomy	−0.0102	0.0030	0.0006	−0.0106	0.989	0.985	0.994	0.896	0.787	1.020
CS	0.0082	0.0006	<0.0001	−0.0026	0.997	0.996	0.999	1.155	0.857	1.555
Uterine procedure	−0.0223	0.0053	<0.0001	−0.0071	0.993	0.988	0.998	0.827	0.626	1.094
Adnexa procedure	−0.0055	0.0031	0.0755	−0.0058	0.994	0.990	0.998	1.079	0.871	1.338
Decline in grade										
Hysterectomy	0.0044	0.0024	0.0633	−0.0049	0.995	0.991	0.999	1.136	1.018	1.267
CS	−0.0153	0.0006	<0.0001	−0.0076	0.992	0.991	0.994	1.251	1.011	1.547
Uterine procedure	0.0008	0.0042	0.8435	−0.0052	0.995	0.990	1.000	1.146	0.921	1.424
Adnexa procedure	−0.0081	0.0030	0.0071	−0.0056	0.994	0.991	0.998	1.233	1.025	1.484
Not changed										
Hysterectomy	−0.0228	0.0012	<0.0001	−0.0078	0.992	0.991	0.994	0.941	0.893	0.992
CS	−0.0038	0.0003	<0.0001	−0.0033	0.997	0.996	0.997	0.831	0.762	0.905
Uterine procedure	0.0000	0.0018	0.9832	−0.0036	0.996	0.995	0.998	0.889	0.813	0.972
Adnexa procedure	−0.0152	0.0013	<0.0001	−0.0011	0.999	0.997	1.001	0.978	0.906	1.055

Adjusted for change in quality assessment of antibiotic use, type of hospital, teaching status, hospital location, number of beds, number of doctors per 100 beds, number of nurses per 100 beds, volume per hospital, CMI, age, CCI, use of laparoscopy and year, 95% CI: confidence interval.

RR (rate ratio): RR is indicated the results of exponentiated estimates and interpretable as percentage changes.

[a]Estimates are the results of log transformation and interpretable as percentage changes.

maximize their performance. Because the quality grades were available to the public during the first year, high-level and unclassified institutions endeavored to maintain or upgrade their quality grade to receive incentive payments; these efforts eventually lead to improvement in quality of care. Conversely, some medical institutions did not make an effort to improve quality of care and experienced decreases in quality and lower condition-specific performance.

In Korea, the VIP and quality assessment have undergone changes as suitable indexes for measuring quality in healthcare have been developed. Quality of care has improved as novel items and expansion of existing clinical conditions have been applied to medical institutes. Quality assessment provides a standardized treatment process that affects physicians' decisions for patient care. As a result, the overall process can be improved for all medical institutions, leading to improved patient outcomes. Quality assessment, however, is dependent on the measurement of quality and methodologies for providing incentives or penalties. Because quality had been assessed for specific medical conditions, this could lead to quality gaps between care of these selected conditions and other conditions in the quality assessment, as well as gaps in quality between the hospitals evaluated in the quality assessment depending on their efforts to improve quality. Some medical institutions may not make an effort to invest in quality improvement, which could result in a decline in quality grade; because the penalty applied to medical institutes with a lower-threshold grade are applied after 2 years. During this 2-year periods, these medical institutions did not receive any intervention from the VIP system, such as compensation and penalty, which may cause deterioration of quality.

In our study lower quality of care was associated with a change in quality assessment grade. The VIP framework considers two aspects, upgrade or maintaining of quality grade; these are considered as an indicator for an incentive or a penalty to medical institutes. This framework does not consider medical institutes with a low performance in patient care. Thus, further refinement of the VIP is needed for improving quality and reducing the quality gap between medical institutes. Policy-makers should consider three aspects for future changes in the VIP. First, multidirectional and comprehensive evaluation of quality of care assessment is needed to achieve the goal of the VIP. Patient outcomes and other performance-related outcomes should be included in this assessment. Second, penalties should be considered not only for lower quality grades, but also for declines in performance. Additional support, such as education programs for medical institutes with low performance, is also needed to reduce the quality gap between medical institutes. Finally, the development of a suitable index that can be used for regular monitoring and to evaluate performance is needed for maintenance of the quality of care.

Our study had several strengths. First, we used NHI claim data, which included large sample sizes of both patients and hospitals. Our results should, therefore, be of significance to policy-makers. Second, to the best of our knowledge, our study was the first to evaluate the relationships between quality assessment and LOS, medical cost and readmission for target and non-target surgeries. Our findings might be also meaningful to other countries with similar VIPs. Third, our results provide valuable evidence to policy-makers regarding the need to modify the VIP aimed at improving quality of care and reducing the quality gap between medical institutes.

Our study also had some limitations. Because we used NHI claim data, we were unable to measure patient socioeconomic status, which is a variable associated with patient outcome. In addition, we could not considered differences within same quality grade in each medical institutes, because, we did not have score of the quality measurement. Finally, due to the limitations of our data set, we could not consider medical institutes that had received incentives or penalties. Further studies are needed to evaluate the effects of incentives or penalties on quality of care at medical institutes.

In conclusion, our results indicate that quality assessment was associated with improvement in quality of care. Our findings provide evidence to policy-makers that quality assessment is important to quality of care and should be developed collaboratively with other quality assessment approaches. Development of counterplans for medical institutes with low quality care is also needed. Regular monitoring and evaluation of the performance of healthcare providers, and additional studies that examine the effects of VIP on quality of care in medical institutes that have received incentives or penalties, are also needed.

References

1. Campbell S, Reeves D, Kontopantelis E *et al*. Quality of primary care in England with the introduction of pay for performance. *N Engl J Med* 2007;**357**:181–90.
2. Petersen LA, Woodard LD, Urech T *et al*. Does pay-for-performance improve the quality of health care? *Ann Intern Med* 2006;**145**:265–72.
3. Lemak CH, Nahra TA, Cohen GR *et al*. Michigan's fee-for-value physician incentive program reduces spending and improves quality in primary care. *Health Aff* 2015;**34**:645–52.
4. Vermeulen MJ, Stukel TA, Boozary AS *et al*. The effect of pay for performance in the emergency department on patient waiting times and quality of care in Ontario, Canada: a difference-in-differences analysis. *Ann Emerg Med* 2016;**67**:496–505.
5. Doran T, Kontopantelis E, Reeves D *et al*. Setting performance targets in pay for performance programmes: what can we learn from QOF. *BMJ* 2014;**348**:g1595.
6. Kim SM, Jang WM, Ahn HA *et al*. Korean National Health Insurance value incentive program: achievements and future directions. *J Prev Med Public Health* 2012;**45**:148.
7. Bisiaux R, Chi Y-L. Republic of Korea: Value Incentive Programme. *Paying for Performance in Health Care Implications for Health System Performance and Accountability: Implications for Health System Performance and Accountability* 2014:251.
8. Kim Y, Kang M. The Performance Management System of the Korean healthcare sector: development, challenges, and future tasks. *Public Perform Manag Rev* 2016;**39**:297–315.
9. Grossbart SR. What's the return? Assessing the effect of "pay-for-performance" initiatives on the quality of care delivery. *Med Care Res Rev* 2006;**63**:29S–48S.
10. Lindenauer PK, Remus D, Roman S *et al*. Public reporting and pay for performance in hospital quality improvement. *N Engl J Med* 2007;**356**:486–96.
11. Van Herck P, De Smedt D, Annemans L *et al*. Systematic review: effects, design choices, and context of pay-for-performance in health care. *BMC Health Serv Res* 2010;**10**:247.
12. Glickman SW, Ou F-S, DeLong ER *et al*. Pay for performance, quality of care, and outcomes in acute myocardial infarction. *JAMA* 2007;**297**:2373–80.
13. Karve AM, Ou F-S, Lytle BL *et al*. Potential unintended financial consequences of pay-for-performance on the quality of care for minority patients. *Am Heart J* 2008;**155**:571–6.
14. McDonald R, Roland M. Pay for performance in primary care in England and California: comparison of unintended consequences. *Ann Fam Med* 2009;**7**:121–7.
15. Sicsic J, Le Vaillant M, Franc C. Building a composite score of general practitioners' intrinsic motivation: a comparison of methods. *Int J Qual Health Care* 2014;**26**:167–73.
16. Scholte M, Neeleman-Van der Steen CW, Hendriks EJ *et al*. Evaluating quality indicators for physical therapy in primary care. *Int J Qual Health Care* 2014;**26**:261–70.
17. Haugum M, Danielsen K, Iversen HH *et al*. The use of data from national and other large-scale user experience surveys in local quality work: a systematic review. *Int J Qual Health Care* 2014;**26**:592–605.
18. Chelmow D, Ruehli MS, Huang E. Prophylactic use of antibiotics for nonlaboring patients undergoing cesarean delivery with intact membranes: a meta-analysis. *Am J Obstet Gynecol* 2001;**184**:656–61.
19. Costantine MM, Rahman M, Ghulmiyah L *et al*. Timing of perioperative antibiotics for cesarean delivery: a metaanalysis. *Am J Obstet Gynecol* 2008;**199**:301. e1–e6.
20. Lee JY, Lee S-I, Jo M-W. Lessons from healthcare providers' attitudes toward pay-for-performance: what should purchasers consider in designing and implementing a successful program? *J Prev Med Public Health* 2012;**45**:137–47.
21. Lee JY, Lee S-I, Kim N-S *et al*. Healthcare organizations' attitudes toward pay-for-performance in Korea. *Health Policy* 2012;**108**:277–85.
22. Kim KH, Park CS, Chang JH *et al*. Association between prophylactic antibiotic use and surgical site infection based on quality assessment data in Korea. *J Prev Med Public Health* 2010;**43**:235–44.
23. Kahn KL, Rogers WH, Rubenstein LV *et al*. Measuring quality of care with explicit process criteria before and after implementation of the DRG-based prospective payment system. *JAMA* 1990;**264**:1969–73.
24. Kahn KL, Rubenstein LV, Draper D *et al*. The effects of the DRG-based prospective payment system on quality of care for hospitalized medicare patients: an introduction to the series. *JAMA* 1990;**264**:1953–5.
25. Feng C, Wang H, Lu N *et al*. Log transformation: application and interpretation in biomedical research. *Stat Med* 2013;**32**:230–9.
26. Faddy M, Graves N, Pettitt A. Modeling length of stay in hospital and other right skewed data: comparison of phase-type, gamma and log-normal distributions. *Value Health* 2009;**12**:309–14.
27. Malehi AS, Pourmotahari F, Angali KA. Statistical models for the analysis of skewed healthcare cost data: a simulation study. *Health Econ Rev* 2015;**5**:1–16.
28. Zeger SL, Liang K-Y, Albert PS. Models for longitudinal data: a generalized estimating equation approach. *Biometrics* 1988:1049–60.
29. Hanley JA, Negassa A, Forrester JE. Statistical analysis of correlated data using generalized estimating equations: an orientation. *Am J Epidemiol* 2003;**157**:364–75.
30. Zou G. A modified poisson regression approach to prospective studies with binary data. *Am J Epidemiol* 2004;**159**:702–6.
31. Jha AK, Orav EJ, Li Z *et al*. The inverse relationship between mortality rates and performance in the Hospital Quality Alliance measures. *Health Aff* 2007;**26**:1104–10.
32. Werner RM, Kolstad JT, Stuart EA *et al*. The effect of pay-for-performance in hospitals: lessons for quality improvement. *Health Aff* 2011;**30**:690–8.