Health Policy and Planning, 2015, 1–8
doi: 10.1093/heapol/czv024
Original article

Does pay-for-performance benefit patients with multiple chronic conditions? Evidence from a universal coverage health care system

Chi-Chen Chen[1] and Shou-Hsia Cheng[2,*]

[1]Department of Public Health, College of Medicine, Fu Jen Catholic University, Taipei, Taiwan and [2]Institute of Health Policy and Management, College of Public Health, National Taiwan University, Taipei, Taiwan

*Corresponding author. Institute of Health Policy and Management, College of Public Health, National Taiwan University, 17, Xu-Zhou Road, Taipei 100, Taiwan. E-mail: shcheng@ntu.edu.tw

Accepted on 11 February 2015

Abstract

Introduction: Numerous studies have examined the impact of pay-for-performance (P4P) programmes, yet little is known regarding their effects on continuity of care (COC) and the role of multiple chronic conditions (MCCs). This study aimed to examine the effects of a P4P programme for diabetes care on health care provision, COC and health care outcomes in diabetic patients with and without comorbid hypertension.

Methods: This study utilized a large-scale natural experiment with a 4-year follow-up period under a compulsory universal health insurance programme in Taiwan. The intervention groups consisted of patients with diabetes who were enrolled in the P4P programme in 2005. The comparison groups were selected via propensity score matching with patients who were seen by the same group of physicians. A difference-in-differences analysis was conducted using generalized estimating equation models to examine the effects of the P4P programme.

Results: Significant impacts were observed after the implementation of the P4P programme for diabetic patients with and without hypertension. The programme increased the number of necessary examinations/tests and improved the COC between patients and their physicians. The programme significantly reduced the likelihood of diabetes-related hospital admissions and emergency department visits [odds ratio (OR): 0.71; 95% confidence interval (CI): 0.63–0.80 for diabetic patients with hypertension; OR: 0.74; 95% CI: 0.64–0.86 for patients without hypertension]. However, the effects of the P4P programme diminished to some extent in the second year after its implementation.

Conclusion: This study suggests that a financial incentive programme may improve the provision of necessary health care, COC and health care outcomes for diabetic patients both with and without comorbid hypertension. Health authorities could develop policies to increase participation in P4P programmes and encourage continued improvement in health care outcomes.

Key words: Continuity of care, diabetes mellitus, health care outcomes, multiple chronic conditions, pay-for-performance programme

Key Messages

- Numerous studies have examined the impact of P4P programmes, yet little is known regarding their effects on COC and the role of MCCs.
- This study demonstrated that enrollment in a diabetes P4P programme led to an increase in the number of necessary examinations/tests, improved the COC between patients and their physicians and reduced the likelihood of diabetes-related hospital admissions and emergency department visits.
- This study also revealed that the previously described effects of the P4P programme were similar for diabetic patients with and without comorbid hypertension.
- This study provides new evidence concerning the beneficial effects of P4P programmes for patients with and without MCCs in a universal coverage health care system.

Introduction

The growing number of patients with multiple chronic conditions (MCCs) has become a challenge for health care systems worldwide. In the United States, over one-quarter of Americans have MCCs (Anderson 2010); care for individuals with MCCs accounts for a disproportionate share of health care expenses (Goodman *et al.* 2012). A study in Canada reported that ~9 of 10 patients aged ≥65 had more than two chronic conditions in general practice settings (Lehnert *et al.* 2011). The development of strategies to manage patients with MCCs is an important task in many industrialized countries.

Pay-for-performance (P4P) is a payment reform that provides financial incentives to health care providers based on specific predetermined quality benchmarks or their provision of proper follow-up care (Petersen *et al.* 2006). A growing literature has examined the effects of P4P programmes, but considerable debate continues regarding their effectiveness on health care outcomes (Petersen *et al.* 2006; Rosenthal and Frank 2006; Christianson *et al.* 2008; Van Herck *et al.* 2010; Emmert *et al.* 2012; Eijkenaar *et al.* 2013). Discrepancies in findings from empirical studies may be attributable to variations in the types of financial incentives implemented, payer mixes and baseline levels of quality of care (Dolor and Schulman 2013). To date, little is known regarding whether P4P programmes for specific diseases have different impacts on health care outcomes in patients with MCCs (Whyte *et al.* 2007; Millett *et al.* 2008; 2009).

Patients with MCCs may benefit from P4P programmes for specific diseases because of the spillover effect. For example, the Physician Group Practice Demonstration Project for diabetes care under the Medicare programme in the United States (Centers for Medicare & Medicaid Services 2011) and the P4P programme for diabetes care in the UK (Doran *et al.* 2006) incorporated blood pressure control as one incentivized outcome measure. A number of studies have demonstrated that patients who were enrolled in P4P programmes for diabetes care were more likely to meet their blood pressure goals (Patric et al. *2006*; Bray *et al.* 2008; Stark *et al.* 2011).

In contrast, a number of other studies have revealed that patients with MCCs may be worse after participating in P4P programmes. First, some studies have argued that evidence-based practice guidelines focus on specific diseases and may not be suitable for providing appropriate care for individuals with MCCs (Tinetti *et al.* 2004; Boyd *et al.* 2005). P4P programmes based on practice guidelines that focus on specific diseases may ignore the varied health care needs of individuals with MCCs (Boyd *et al.* 2005). Second, P4P programmes may be harmful for individuals with MCCs because health care providers may 'cherry pick' their patients. Previous studies have indicated that older patients and patients with more comorbidities or more severe conditions are more likely to be excluded from P4P programmes (Shen 2003; Chen *et al.* 2011; Chang *et al.* 2012). Third, a previous study in the UK also demonstrated that P4P programmes can hamper the continuity of care (COC) between patients and their physicians (Campbell *et al.* 2009). In addition, fragmentation in care is commonly seen among patients with MCCs, which implies that P4P programmes may decrease the COC for these individuals.

The health care system in Taiwan

Taiwan's health care system focuses on specialist and hospital care, and there is no referral requirement or gatekeeper in primary care settings. Therefore, patients are free to visit physicians at community clinics or hospital outpatient departments for any episode based on their preference. In addition, Taiwan's National Health Insurance (NHI) programme was launched in 1995 and has improved public access to health care services (Cheng 2003). As a result of the freedom of choice and the easy access to ambulatory care, the average number of annual western physician visits was ~13 per capita in 2011 (National Health Insurance Administration 2014). Accordingly, patients are often criticized for their doctor shopping behaviour, which may hamper the COC between patients and their physicians (Chen *et al.* 2006).

P4P programmes in Taiwan

Since 2001, the Bureau of the NHI (renamed the NHI Administration in 2013) has implemented a P4P programme for diabetes care; physicians who are metabolic or endocrinology specialists or physicians who have participated in a training programme for diabetes shared care can voluntarily apply to participate in the programme. Participating physicians can then invite individual patients to enroll. This programme intended to promote guideline-based practice with financial incentives under a fee-for-services payment scheme. In addition to regular reimbursement for health care services such as physician visits, medications, physical examinations and laboratory tests, the P4P programme compensates participating clinicians' additional 'enlarged physician fees' and 'case management fees.' Required and recommended services included in these initial and follow-up visits (e.g. diabetes-specific eye examination, laboratory evaluation and self-care education) are clearly defined by the P4P programme. More detailed information concerning the design of the P4P diabetes programme's payment incentives can be found elsewhere (Cheng *et al.* 2012).

Previous studies examining the effects of the P4P programmes have found that the P4P programme for diabetes care has resulted in quality improvement regarding health care provision and health care outcomes (Lee *et al.* 2010; Cheng *et al.* 2012) and reduction in

overall health care expenses in the long run (Cheng *et al.* 2012). Yet, other studies have revealed unintended effects of the P4P programme such as 'cherry picking' of healthier patients (Chen *et al.* 2011; 2012) and higher risk of emergency department (ED) visits due to diabetic hypoglycemia (Yu *et al.* 2014).

This study extends the existing literature in two ways. First, the majority of evidence regarding the impact of P4P programmes on health care outcomes reflects average population effects, but little is known concerning their impact on patients with MCCs. Second, the majority of previous studies focus on the effects of P4P programmes on health care utilization and outcomes, as well as expenses (Petersen *et al.* 2006). However, there is limited evidence regarding their effects on COC (Van Herck *et al.* 2010). Because hypertension is a common comorbid condition for diabetic patients (Epstein and Sowers 1992), this study aims to examine the effects of a P4P programme for diabetes care on health care provision, COC and health care outcomes in diabetic patients with and without comorbid hypertension.

Materials and methods

Data source and study sample

This study employed a natural experiment design and was based on a 4-year panel of claims data on health care utilization, from 2004 to 2007, which was provided by the National Health Research Institute in Taiwan. This dataset contains detailed records of every physician visit and hospital admission for each patient, such as primary and secondary diagnosis codes, physical examination and laboratory tests codes. Patients with a diagnosis of type 2 diabetes were identified using the International Classification of Diseases, 9th revision-Clinical Modification (ICD-9-CM) codes 250.xx (excluding type 1 diabetes codes 250.x1 or 250.x3) and had at least three diabetes-related physician visits or at least one diabetes-related hospitalization each year from 2004 to 2007. Patients < 18 years of age were excluded.

Patients with diabetes who were enrolled in the diabetes P4P programme in 2005 were defined as the intervention group. The index date for each patient was defined as the date each patient was first enrolled in the programme between 1 January 2005 and 31 December 2005. Because the patients who participated in the programme in 2005 might have existed in subsequent years, this study included only the subjects who remained in the programme (with specific P4P claims) every year from 2005 to 2007. As a result, 8351 patients comprised the intervention group, and the potential comparison group consisted of patients with diabetes who had never been enrolled in the P4P programme between 2001 and 2007($n = 178\,892$).

Because the patients who were enrolled in the programme were purposively selected by their physicians, a potential selection bias could have hampered the comparability between the intervention and comparison groups. In this study, we used two strategies to increase the comparability between intervention and comparison groups; one strategy was to select patients from the same group of physicians, and the other strategy was to use propensity score matching (PSM). First, we identified the most frequently visited physician of the 8351 patients who were enrolled in the P4P programme. Then, all diabetic patients who visited the same group of physicians were identified, and the patients who had never been enrolled in the P4P programme were considered potential candidates for the comparison group. There were 2003 frequently visited physicians of the 8351 patients in the intervention group and 47 214 potential candidates in the comparison group.

Second, we used a PSM approach to select patients with diabetes for the comparison group to minimize selection bias (Rosenbaum and Rubin 1983). We created a propensity score for each patient that estimated the probability of enrolment in the P4P programme based on the subject's characteristics using a generalized estimating equation (GEE) model with binary distribution and logit link. We used this model to account for the effects of patient clustering among particular physicians (Fitzmaurice *et al.* 2004). The characteristics included each patient's age, sex, hypertension status, diabetes complication severity index (DCSI) score (Young *et al.* 2008) and chronic illness with complexity (CIC) index score (Meduru *et al.* 2007), in addition to the location and accreditation level of each patient's most frequently visited health care provider. The four accreditation levels (in descending order) were medical centre, regional hospital, district hospital and community clinic (Huang *et al.* 2000). The DCSI consists of scores (no abnormality = 0, some abnormality = 1 or severe abnormality = 2) in seven categories of complications: cardiovascular complications, nephropathy, retinopathy, peripheral vascular disease, stroke, neuropathy and metabolic disorders. Neuropathy is coded as 0 or 1 only. The highest score on the DCSI is 13 (Young *et al.* 2008). The CIC index was used to adjust for comorbidity in patients with MCCs. This index contains information regarding non-diabetes physical illness complexity (including cancers gastrointestinal, musculoskeletal and pulmonary diseases), diabetes-related complexity and mental illness/substance abuse complexity (including substance abuse and mental illness) (Meduru *et al.* 2007). We excluded diabetes-related complexity to avoid collinearity with the comorbidity effect captured by the DCSI score. In the analysis, the CIC index was calculated as the sum of the previously described categories, which ranged from 0 to 6.

We employed the caliper matching method with 1:2 matching between the intervention and comparison groups based on their propensity scores. For the 8351 patients who participated in the P4P programme, the PSM yielded 16 702 patients in the comparison group. The study period ranged from 1 year before the index date to 2 years of subsequent follow up. Because the patients in the comparison group lacked index dates for enrolment in the P4P programme, their study periods were determined by the index date of their matched counterparts in the intervention group. As a result, 25 053 patients and 75 159 patient-years were included in the analysis. The unit of analysis was patient-years.

In terms of MCCs, we selected diabetes and hypertension to represent the patients' MCC status because these two diagnoses are commonly seen together (Epstein and Sowers 1992). Diabetes mellitus was considered the index condition, and hypertension status (ICD-9-CM codes 401–405) was treated as a comorbid condition in this study. We defined an individual as having hypertension if they had at least three hypertension-related physician visits in the claims records each year. The patients were categorized into two mutually exclusive groups: the diabetic patients with comorbid hypertension and the diabetic patients without comorbid hypertension. The patient-years were then calculated for each group, which accounted for the dynamic changes in comorbid hypertension status over the study period.

Measures of study variables

Dependent variable

Three outcome measures were examined in this study: the number of essential examinations/tests the patients received, COC and health care outcomes. The number of essential examination/tests included ophthalmoscopic examinations and the following

laboratory tests: blood glucose, HbA1c, lipid profile, serum creatinine, SGPT/ALT and urinalysis.

COC measures were classified into five types, i.e. duration, density, dispersion, sequence and subjective measures (Jee and Cabana 2006). When using a claims dataset for analyses, researchers tend to use indices such as the usual provider of care (UPC) index to measure the density of visiting a physician frequently, the COC index (COCI) to measure the dispersion of visits and the sequential continuity (SECON) index to sequentially measure the various physicians visited. Previous studies have reported that, compared with the UPC and SECON indices, the COCI is less sensitive to the number of physician visits (Smedby 1986). Considering the variation and very high number of physician visits in Taiwan, we chose the COCI as our primary dependent variable.

The COCI score was based on the number of different physicians seen and the number of visits to each physician each year during the study period. The equation for this index is as follows:

$$\mathrm{COCI} = \frac{\sum_{j=1}^{M} n_j^2 - N}{N(N-1)},$$

where N represents the total number of physician visits, n_j is the number of visits to the same physician, j represents a given physician and M is the total number of physicians. In a study by Bice and Boxerman (1977), the summation term in the numerator was the sum of the number of un-referred physicians. Because of the lack of referral arrangements in Taiwan, we used the total number of physicians in the analysis. To increase the comparability of the COC among patients, we excluded the following treatment categories when determining the total number of physician visits: outpatient surgery, dental care and specific services, such as long-term care. This index measures the degree to which patient visits are dispersed among different physicians (from 0 to 1); a higher value corresponds to better COC. Because the COC values have no inherent clinical meaning, we categorized the values into two groups based on the observations that were below or above the median COCI score. Finally, the health care outcome measure was whether the patient was hospitalized or had an ED visit for diabetes-related conditions during each year of the study period.

Independent variable and covariates

The main independent variables were the patients' enrollment in the P4P programme, the time dummy variables for the 2 years after the index date and the two interaction terms for the previously described variables. A number of covariates were controlled for in the regression models: the patient's characteristics (age, sex, DCSI and CIC index scores), as well as the provider's characteristics (accreditation level and location).

Statistical analysis

This study used a difference-in-differences (DID) analysis (also called a pre- and post-design with a comparison group) to compare the outcomes between the two groups before and after the implementation of the P4P programme. GEE models were used to account for the intraclass correlation between repeated observations for the same patients and patients in the same matched pairs (Fitzmaurice *et al.* 2004). Based on the variables under investigation, the number of examinations/tests was analyzed using a Poisson distribution and a logarithmic link function, and the COC status (high or low) and the likelihood of hospitalizations or ED visits were analysed using a binominal distribution and logit link function. The analyses were performed using SAS version 9.1.3 (SAS Institute, Cary, NC) and STATA 9.1 (Stata Corp., College Station, TX).

In addition to these estimation models, we conducted two sensitivity analyses to verify the robustness of the findings. First, we employed the UPC index to measure the COC. The UPC index was defined as the number of physician visits to the most frequently seen physician divided by the total number of physician visits. Second, we used a continuous scale for the COCI and the UPC index in the analyses.

Results

Table 1 shows the baseline characteristics of the patients in the intervention and comparison groups in the pre-matched and post-matched samples. For the pre-matched sample, significant differences were detected between the intervention and comparison groups ($P < 0.001$). The PSM process resulted in a more balanced distribution of the patient and hospital characteristics between the two study groups.

Descriptions of the study variables for each year during the study period are presented in Table 2. Similar results were observed for the diabetic patients with and without comorbid hypertension. The average number of essential examinations/tests performed was similar between the two study groups in the baseline year (4.08 vs 3.94 for diabetic patients with comorbid hypertension; 3.84 vs 3.76 for diabetic patients without comorbid hypertension). After the first year of the P4P programme, the average number of examinations/tests performed increased in the intervention group for the diabetic patients with hypertension (4.08–6.37) and for those without hypertension (3.84–6.37). However, the figures declined slightly in the second year of the P4P programme. In contrast, the number of examinations /tests for the patients in the comparison group increased slightly and stably over the same period of time.

The COCI values for the diabetic patients without comorbid hypertension were slightly higher than the scores for the patients with comorbid hypertension in both the intervention and comparison groups. It appears that enrolment in the P4P programme increased the COCI scores in the first year but that the scores decreased slightly in the second year; the scores were 0.31, 0.34 and 0.33 for the diabetic patients with hypertension and 0.33, 0.36 and 0.35 for the diabetic patients without hypertension. The COCI scores remained stable during this period in the comparison group; the figures were 0.33, 0.34 and 0.33 for the patients with hypertension and 0.35, 0.36 and 0.37 for the patients without hypertension.

Finally, the rates of hospital admissions or ED visits for diabetes-related conditions were higher in the group with comorbid hypertension. The rates steadily increased in the comparison group over the 3 years; the figures were 22.79, 23.88 and 26.20% in the diabetic patients with hypertension and 18.95, 19.05 and 20.81% in the diabetic patients without hypertension. However, there were clear changes in the intervention group after the patients were enrolled in the P4P programme; the figures were 23.97, 20.40 and 22.53% in the patients with hypertension and 18.00, 15.16 and 17.61% in the patients without hypertension.

Table 3 presents the DID analysis results for the effects of the diabetes P4P programme with GEE models by comorbid hypertension status. With regard to the number of examinations/tests, the DID parameter, the coefficients of the interaction terms, was positive and significant in the first and second years after the intervention ($\beta = 0.430$ and 0.309, respectively; $P < 0.001$) in the diabetic patients with hypertension; similar effects were observed in the patients without hypertension ($\beta = 0.475$ and 0.348, respectively; both $P < 0.001$). Regarding the effects of the P4P programme on the patients' COC status, we demonstrated that the DID parameter was

Table 1. Characteristics of the patients with diabetes in the pre-matched and post-matched samples in the baseline year

Characteristics	Pre-matched sample					Post-matched sample				
	Control group		Intervention group		*P* value	Control group		Intervention group		*P* value
	N	%	*N*	%		*N*	%	*N*	%	
Total	**47 214**		**8351**			**16 702**		**8351**		
Sex					<0.001					0.516
Female	25 772	54.59	4762	57.02		9452	56.59	4762	57.02	
Male	21 442	45.41	3589	42.98		7250	43.41	3589	42.98	
Age groups (*N*, %)					<0.001					0.984
<55	11 207	23.74	2044	24.48		4080	24.43	2,044	24.48	
56–65	13 251	28.07	2711	32.46		5441	32.58	2,711	32.46	
66–75	14 982	31.73	2672	32.00		5333	31.93	2672	32.00	
76+	7774	16.47	924	11.06		1848	11.06	924	11.06	
Hypertension	28 716	60.82	5163	61.82	0.083	10 418	62.38	5163	61.82	0.397
DCSI score					<0.001					0.472
Score 0	17 566	37.21	2505	30.00		5119	30.65	2505	30.00	
Score 1	13 761	29.15	2640	31.61		5288	31.66	2640	31.61	
Score 2+	15 887	33.65	3206	38.39		6295	37.69	3206	38.39	
CIC index score					<0.001					0.986
Score 0	16 699	35.37	2774	33.22		5542	33.18	2774	33.22	
Score 1	18 169	38.48	3205	38.38		6399	38.31	3205	38.38	
Score 2+	12 346	26.15	2372	28.40		4761	28.51	2372	28.40	
Accreditation level of hospital					<0.001					0.538
Medical centre	17 729	37.55	1748	20.93		3495	20.93	1748	20.93	
Regional hospital	10 349	21.92	2003	23.99		4141	24.79	2003	23.99	
District hospital	6192	13.11	1667	19.96		3272	19.59	1667	19.96	
Community clinic	12 944	27.42	2933	35.12		5794	34.69	2933	35.12	
Location of hospital					<0.001					0.910
Taipei and northern region	23 376	49.51	4040	48.38		8106	48.53	4040	48.38	
Central and southern region	13 780	29.19	3284	39.32		6523	39.06	3284	39.32	
Kao-ping and eastern region	10 058	21.30	1027	12.30		2073	12.41	1027	12.30	

DCSI score, diabetes complication severity index score; CIC index score, chronic illness with complexity index score.

Table 2. Distribution of research variables by year and comorbid hypertension status

Variables	Diabetes with hypertension						Diabetes without hypertension					
	Pre-P4P		Year 1		Year 2		Pre-P4P		Year 1		Year 2	
Number of essential examinations/tests (mean, SD)												
Intervention group	4.08	1.52	6.37	1.00	5.81	1.28	3.84	1.59	6.37	1.00	5.80	1.33
Comparison group	3.94	1.64	4.00	1.62	4.13	1.58	3.76	1.74	3.85	1.71	3.98	1.65
COCI (mean, SD)												
Intervention group	0.31	0.21	0.34	0.21	0.33	0.21	0.33	0.22	0.36	0.22	0.35	0.22
Comparison group	0.33	0.23	0.34	0.23	0.33	0.22	0.35	0.23	0.36	0.24	0.37	0.24
Diabetes-related hospitalizations or ED visits (*N*, %)												
Intervention group	1122	23.97	992	20.40	1146	22.53	625	18.00	499	15.16	540	17.61
Comparison group	2194	22.79	2379	23.88	2673	26.20	1266	18.95	1208	19.05	1270	20.81

COCI, continuity of care index; ED visits, emergency department visits.

also positive and significant, with an odds ratio (OR) = 1.42 and a 95% confidence interval (CI) = 1.30–1.54 in the first year after the intervention and an OR = 1.26 with a 95% CI = 1.15–1.36 in the second year for the diabetic patients with hypertension. For the patients without hypertension, the figures were similar, with an OR = 1.46 and 1.20, respectively, and a 95% CI = 1.32–1.61 and 1.09–1.33, respectively.

Finally, the net effect (DID parameter) of the P4P programme on the likelihood of hospital admissions or ED visits was negative and significant, with an OR = 0.71 and a 95% CI = 0.63–0.80 in the first year after the intervention and an OR = 0.73 and a 95% CI = 0.65–0.81 in the second year for the diabetic patients with hypertension. Similar results were also found for the patients without hypertension, with an OR = 0.74 and a 95% CI = 0.64–0.86 in the first year after the intervention and an OR = 0.80 and a 95% CI = 0.69–0.93 in the second year.

Sensitivity analyses

We conducted two sensitivity analyses to examine the robustness of the findings. First, using the UPC index as the COC indicator, we determined that the results were similar to the results obtained using the COCI (Supplementary table). Second, using a continuous scale instead of a dichotomous measure (high or low) for the COCI and the UPC index, we identified similar results (Supplementary table).

Table 3. Adjusted GEE estimations of the effects of P4P programme on the outcome variables

Variables	Number of essential examinations/tests				COC index						DM-related hospitalizations or ED visits					
	With hypertension		Without hypertension		With hypertension			Without hypertension			With hypertension			Without hypertension		
	β	*P* value	β	*P* value	OR	95% CI		OR	95% CI		OR	95% CI		OR	95% CI	
Intervention group	0.043	0.027	0.040	0.086	0.79	0.73	0.85	0.77	0.71	0.84	1.09	1.00	1.19	0.98	0.88	1.10
Period (Reference group: pre-P4P)																
Post-P4P Year1	0.017	0.193	0.020	0.199	1.04	0.99	1.09	1.07	1.01	1.13	1.06	0.99	1.13	0.97	0.89	1.06
Post-P4P Year2	0.051	<0.001	0.054	0.001	1.01	0.96	1.06	1.07	1.01	1.14	1.20	1.12	1.28	1.06	0.98	1.16
Interaction term between P4P and period																
Intervention group*Post-P4P Year 1	0.430	<0.001	0.475	<0.001	1.42	1.30	1.54	1.46	1.32	1.61	0.71	0.63	0.80	0.74	0.64	0.86
Intervention group*Post-P4P Year 2	0.309	<0.001	0.348	<0.001	1.26	1.15	1.36	1.20	1.09	1.33	0.73	0.65	0.81	0.80	0.69	0.93
Male	0.006	0.661	0.003	0.878	1.15	1.09	1.21	1.26	1.19	1.34	0.89	0.84	0.94	0.89	0.83	0.95
Age group (Reference group: <55)																
56–65	−0.003	0.867	−0.004	0.851	0.93	0.87	1.00	0.93	0.86	1.00	1.05	0.97	1.14	1.03	0.94	1.12
66–75	−0.030	0.116	−0.023	0.277	0.91	0.85	0.98	0.90	0.84	0.98	1.21	1.11	1.31	1.25	1.14	1.37
76+	−0.071	0.002	−0.065	0.025	0.95	0.87	1.03	0.94	0.84	1.04	1.48	1.35	1.62	1.70	1.51	1.91
DCSI score (Reference group: Score 0)																
Score 1	0.057	<0.001	0.076	<0.001	0.73	0.70	0.77	0.80	0.75	0.85	1.46	1.36	1.57	1.29	1.18	1.41
Score 2+	0.111	<0.001	0.111	<0.001	0.44	0.42	0.46	0.51	0.48	0.54	3.52	3.30	3.75	2.86	2.64	3.10
CIC index score (Reference group: Score 0)																
Score 1	0.040	0.001	0.058	<0.001	0.50	0.48	0.52	0.56	0.53	0.59	1.65	1.56	1.75	1.54	1.43	1.67
Score 2+	0.079	<0.001	0.117	<0.001	0.26	0.25	0.28	0.30	0.28	0.32	2.62	2.46	2.79	2.47	2.27	2.69
Accreditation level of hospital (Reference group: community clinic)																
Medical centre	0.167	<0.001	0.214	<0.001	0.50	0.47	0.54	0.65	0.60	0.70	1.44	1.34	1.55	1.55	1.41	1.71
Regional hospital	0.146	<0.001	0.183	<0.001	0.60	0.56	0.63	0.69	0.64	0.74	1.90	1.78	2.03	1.81	1.65	1.97
District hospital	0.112	<0.001	0.151	<0.001	0.65	0.61	0.69	0.66	0.61	0.72	1.47	1.36	1.57	1.47	1.33	1.62
Location of hospital (Reference group: Kao-ping and eastern region)																
Taipei and northern region	0.085	<0.001	0.084	0.002	1.31	1.21	1.42	1.31	1.18	1.44	0.91	0.84	0.99	0.77	0.69	0.86
Central and southern region	0.017	0.422	0.009	0.733	1.27	1.17	1.37	1.30	1.18	1.44	0.96	0.89	1.04	0.88	0.79	0.98

GEE, generalized estimating equation; P4P programme, pay-for-performance programme; DCSI score, diabetes complication severity index score; CIC index score, chronic illness with complexity index score; OR, odds ratio; 95% CI, 95% confidence interval.

The findings from the sensitivity analyses revealed that the effects of the P4P programme were robust in this study.

Discussion

The main purpose of this study was to examine the effects of a diabetes P4P programme on health care provision, COC and health care outcomes in diabetic patients with and without comorbid hypertension. The results indicated that the diabetes P4P programme led to an increase in the number of necessary examinations/tests and improved the COC between patients and their physicians. The results also indicated that the programme significantly reduced the patients' likelihood of hospital admissions or ED visits. Furthermore, the effects were similar for diabetic patients with and without comorbid hypertension. However, these programme effects diminished to some extent in the second year after its implementation.

This study demonstrated that patients enrolled in the diabetes P4P programme received more essential examinations/tests than did their counterparts; this finding is similar to previous studies that used a natural experiment design (Beaulieu and Horrigan 2005; Levin-Scherz *et al.* 2006). In addition, we demonstrated that patients in the P4P programme had a lower probability of diabetes-related hospital admissions or ED visits. This finding is also consistent with previous reports (Lee *et al.* 2010; Cheng *et al.* 2012). The decreased diabetes-related hospitalizations and ED visits might reflect better quality of ambulatory care under the P4P programme. For example, the programme could have increased physicians' intention to follow up on glycemic status and detect minor conditions at earlier stages. It is noteworthy that the effects of the programme diminished in the second year after its introduction. Two recent studies also revealed that the short-term improvements in quality measures with respect to incentives may not be sustained (Campbell *et al.* 2009; Kristensen *et al.* 2014). We suggest that continuous improvement is required and could be achieved by modifying the incentive design or performing more detailed monitoring of the programme's performance.

There is a large body of studies concerning the effects of P4P programmes on health care utilization and outcomes, but limited evidence has been reported for the effects of these programmes on COC (Van Herck *et al.* 2010). The findings from this study demonstrated that the COC levels increased in diabetic patients after their enrollment in the P4P programme compared with the patients who did not enroll. This finding is different from that reported by Campbell *et al.* (2009); they used an interrupted time-series analysis and reported that patients' COC declined after the introduction of a P4P programme in the UK. One potential explanation for the inconsistent results is the characteristics of the health care systems in Taiwan and the UK. In the UK, individuals are required to follow a formal referral arrangement coordinated by primary care gatekeepers, and the enhancement of access to health care services is an important task for health authorities. One of the purposes of the UK P4P programme was to increase rapid access to primary care (within 48 h), which could have impeded the COC between patients and their family doctors (Campbell *et al.* 2009). Conversely, patients in Taiwan can choose preferred physicians for any episode without referral; the easy access to health care might lead to doctor shopping (Chen *et al.* 2006). We considered that the diabetes P4P programme improved the provision of necessary examinations/tests and comprehensive follow-up care for patients, which could have improved patient–physician COC via the increase of mutual trust and the enhancement of information sharing between patients and their physicians.

One objective of this study was to examine the effects of a P4P programme on health outcomes in diabetic patients with and without comorbid hypertension. The majority of previous studies have only focused on a single chronic disease (Petersen *et al.* 2006); only a limited number of studies have examined the effects of P4P programmes for patients with comorbid conditions, and the results have been mixed. Two studies demonstrated that patients with comorbid conditions appeared to have benefited more from the P4P programme than did those without comorbid conditions (Millett *et al.* 2008; 2009). Conversely, one study demonstrated that the overall quality of diabetes management was similar for patients with and without schizophrenia or bipolar disorder in the first year of a new contract for the General Medical Services programme (Whyte *et al.* 2007). In this study, we found that the beneficial effects of the P4P programme, including the provision of necessary health care follow-ups, COC and health care outcomes, were similar in diabetic patients with and without hypertension. Our findings suggest that P4P programmes may also benefit patients with comorbid conditions. Unfortunately, previous studies have found that these programmes might lead to unintended consequences, such as 'cherry picking' of healthier patients (Chen *et al.* 2010; Chang *et al.* 2012). For example, Chen *et al.* (2010) found that older patients and patients with more comorbidities or more severe conditions were more likely to be excluded from P4P programmes. In addition, the participation rate in P4P programmes for diabetes care remains low (30%) in Taiwan (National Health Insurance Administration 2014). Therefore, improvements in the participation rate in P4P programmes or the introduction of compulsory participation for patients with MCCs should be pursued.

This study has a number of limitations. First, we did not include certain unobserved (such as health literacy) or unavailable (such as education level) characteristics in the PSM GEE models. These uncontrolled characteristics could have contributed to the pre-existing differences between the intervention and comparison groups. Nevertheless, this concern might have been mitigated because we employed a DID analysis with a longitudinal study design. Second, there is no consensus on the definition and measurement of MCCs. In this study, we only included hypertension as the comorbid condition, which could be far from as comprehensive as one might expect. Third, there are certain unique features of Taiwan's health care system that may limit the generalizability of the findings to other populations.

Despite the limitation, this study provides evidence that a financial incentive programme appears to be successful in improving the provision of necessary health care, COC and health care outcomes in Taiwan. The beneficial effects of the P4P programme were similar for diabetic patients with and without comorbid hypertension. More attention should be directed towards the development of new policies to improve the participation rate in P4P programmes and to encourage continued improvements in health outcomes. Moreover, it is recommended to investigate the reasons for inconsistent findings concerning the impacts of P4P programmes in different health care systems for future studies.

Supplementary Data

Supplementary data are available at *HEAPOL* online.

Funding

The study was supported by the grants from the Ministry of Science and Technology (MOST 100-2410-H-002-084 and MOST 103-2410-H-002-202) and the National Health Research Institute (NHRI-EX104-10225PI) in Taiwan. The funding source had no role in the study.

Conflict of interest

None declared.

References

Anderson G. 2010. *Chronic Conditions: Making the Case for Ongoing Care*. Princeton, NJ: Robert Wood Johnson Foundation and John Hopkins University.

Beaulieu ND, Horrigan DR. 2005. Putting smart money to work for quality improvement. *Health Services Research* **40**: 1318–34.

Bice TW, Boxerman SB. 1977. A quantitative measure of continuity of care. *Medical Care* **15**: 347–9.

Boyd CM, Darer J, Boult C *et al*. 2005. Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. *The Journal of the American Medical Association* **294**: 716–24.

Bray K, Turpin RS, Jungkind K, Heuser G. 2008. Defining success in diabetes disease management: digging deeper in the data. *Disease Management* **11**: 119–28.

Campbell SM, Reeves D, Kontopantelis E, Sibbald B, Roland M. 2009. Effects of pay for performance on the quality of primary care in England. *The New England Journal of Medicine* **361**: 368–78.

Centers for Medicare & Medicaid Services. Medicare physician group practice demonstration. https://www.cms.gov/Medicare/Demonstration-Projects/DemoProjectsEvalRpts/downloads/PGP_Fact_Sheet.pdf, accessed 21 August 2014.

Chang RE, Lin SP, Aron DC. 2012. A pay-for-performance program in Taiwan improved care for some diabetes patients, but doctors may have excluded sicker ones. *Health Affairs (Millwood)* **31**: 93–102.

Chen TJ, Chou LF, Hwang SJ. 2006. Patterns of ambulatory care utilization in Taiwan. *BMC Health Services Research* **6**: 54.

Chen TT, Chung KP, Lin IC, Lai MS. 2011. The unintended consequence of diabetes mellitus pay-for-performance (P4P) program in Taiwan: are patients with more comorbidities or more severe conditions likely to be excluded from the P4P program? *Health Services Research* **46**: 47–60.

Cheng SH, Lee TT, Chen CC. 2012. A longitudinal examination of a pay-for-performance program for diabetes care: evidence from a natural experiment. *Medical Care* **50**: 109–16.

Cheng TM. 2003. Taiwan's new national health insurance program: genesis and experience so far. *Health Affairs (Millwood)* **22**: 61–76.

Christianson JB, Leatherman S, Sutherland K. 2008. Lessons from evaluations of purchaser pay-for-performance programs: a review of the evidence. *Medical Care Research and Review* **65**: 5S–35S.

Dolor RJ, Schulman KA. 2013. Financial incentives in primary care practice: the struggle to achieve population health goals. *The Journal of the American Medical Association* **310**: 1031–2.

Doran T, Fullwood C, Gravelle H, *et al*. 2006. Pay-for-performance programs in family practices in the United Kingdom. *The New England Journal of Medicine* **355**: 375–84.

Eijkenaar F, Emmert M, Scheppach M, Schöffski O. 2013. Effects of pay for performance in health care: a systematic review of systematic reviews. *Health Policy* **110**: 115–30.

Emmert M, Eijkenaar F, Kemter H, Esslinger AS, Schöffski O. 2012. Economic evaluation of pay-for-performance in health care: a systematic review. *The European Journal of Health Economics* **13**: 755–67.

Epstein M, Sowers JR. 1992. Diabetes mellitus and hypertension. *Hypertension* **19**: 403–18.

Fitzmaurice GM, Laird NM, Ware JH. 2004. *Applied Longitudinal Analysis*. Hoboken, NJ: Wiley.

Goodman RA, Parekh AK, Koh HK. 2012. Toward a more cogent approach to the challenges of multimorbidity. *Annals of Family Medicine* **10**: 100–1.

Huang P, Hsu YH, Kai-Yuan T, Hsueh YS. 2000. Can European external peer review techniques be introduced and adopted into Taiwan's hospital accreditation system? *International Journal for Quality in Health Care* **12**: 251–4.

Jee SH, Cabana MD. 2006. Indices for continuity of care: a systematic review of the literature. *Medical Care Research and Review* **63**: 158–88.

Kristensen SR, Meacock R, Turner AJ, *et al*. 2014. Long-term effect of hospital pay for performance on mortality in England. *The New England Journal of Medicine* **371**: 540–8.

Lee TT, Cheng SH, Chen CC, Lai MS. 2010. A pay-for-performance program for diabetes care in Taiwan: a preliminary assessment. *The American Journal of Managed Care* **16**: 65–9.

Lehnert T, Heider D, Leicht H, *et al*. 2011. Review: health care utilization and costs of elderly persons with multiple chronic conditions. *Medical Care Research and Review* **68**: 387–420.

Levin-Scherz J, DeVita N, Timbie J. 2006. Impact of pay-for-performance contracts and network registry on diabetes and asthma HEDIS measures in an integrated delivery network. *Medical Care Research and Review* **63**: 14S–28S.

Meduru P, Helmer D, Rajan M, *et al*. 2007. Chronic illness with complexity: implications for performance measurement of optimal glycemic control. *Journal of General Internal Medicine* **22**: 408–18.

Millett C, Gray J, Bottle A, Majeed A. 2008. Ethnic disparities in blood pressure management in patients with hypertension after the introduction of pay for performance. *Annals of Family Medicine* **6**: 490–6.

Millett C, Bottle A, Ng A, *et al*. 2009. Pay for performance and the quality of diabetes management in individuals with and without co-morbid medical conditions. *Journal of the Royal Society of Medicine* **102**: 369–77.

National Health Insurance Administration. 2014. Statistical trends of National Health Insurance. http://www.nhi.gov.tw/Resource/webdata/23838_1_%E5%85%A8%E6%B0%91%E5%81%A5%E5%BA%B7%E4%BF%9D%E9%9A%AA%E7%B5%B1%E8%A8%88%E5%8B%95%E5%90%91- 2011%E5 % B9%B4.pdf, accessed 19 December 2014.

National Health Insurance Administration. 2014. The participation rate of P4P programs. http://www.nhi.gov.tw/webdata/webdata.aspx?menu=17&menu_id=659&WD_ID=897&webdata_id=4531, accessed 21 August 2014.

Patric K, Stickles JD, Turpin RS *et al*. 2006. Diabetes disease management in Medicaid managed care: a program evaluation. *Disease Management* **9**: 144–56.

Petersen LA, Woodard LD, Urech T, Daw C, Sookanan S. 2006. Does pay-for-performance improve the quality of health care? *Annals of Internal Medicine* **145**: 265–72.

Rosenbaum PR, Rubin DB. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* **70**: 41–55.

Rosenthal MB, Frank RG. 2006. What is the empirical basis for paying for quality in health care? *Medical Care Research and Review* **63**: 135–57.

Shen Y. 2003. Selection incentives in a performance-based contracting system. *Health Services Research* **38**: 535–52.

Smedby O, Eklund G, Eriksson EA, Smedby B. 1986. Measures of continuity of care. A register-based correlation study. *Medical Care* **24**: 511–8.

Stark RG, Schunk MV, Meisinger C *et al*. 2011. Medical care of type 2 diabetes in German disease management programmes: a population-based evaluation. *Diabetes/Metabolism Research and Reviews* **27**: 383–91.

Tinetti ME, Bogardus ST Jr, Agostini JV. 2004. Potential pitfalls of disease-specific guidelines for patients with multiple conditions. *The New England Journal of Medicine* **351**: 2870–4.

Van Herck P, De Smedt D, Annemans L *et al*. 2010. Systematic review: effects, design choices, and context of pay-for-performance in health care. *BMC Health Services Research* **10**: 247.

Whyte S, Penny C, Phelan M, Hippisley-Cox J, Majeed A. 2007. Quality of diabetes care in patients with schizophrenia and bipolar disorder: cross-sectional study. *Diabetic Medicine* **24**: 1442–8.

Young BA, Lin E, Von Korff M *et al*. 2008. Diabetes complications severity index and risk of mortality, hospitalization, and healthcare utilization. *The American Journal of Managed Care* **14**: 15–23.

Yu HC, Tsai WC, Kung PT. 2014. Does the pay-for-performance programme reduce the emergency department visits for hypoglycaemia in type 2 diabetic patients? *Health Policy and Planning* **29**: 732–41.