ô Open Access Full Text Article

ORIGINAL RESEARCH

Changes in hospital performance after implementation of pay-for-performance in Japan

Yusuke Inoue
Seungwon Jeong
Katsunori Kondo

Nihon Fukushi University, Chiyoda, Nakaku, Nagoya, Japan

Background: The purpose of this study was to evaluate the effect of pay-for-performance (P4P) in hospitals in Japan, and to determine if any improvement occurred in the Functional Independence Measure (FIM™) score at the time of discharge from hospital, the return home rate, and whether there was any intentional selection by hospitals of patients expected to have greater recovery.

Methods: We used the data produced by the Rehabilitation Patients DataBank of Japan (February 2011, for 903 patients and 26 hospitals), and identified patient characteristics and processes (by *t*-test and Chi-square test), change in FIM score on discharge from hospital (by regression analysis), and change in return home rate (by logistic regression analysis) before and after the implementation of P4P.

Results: Given the high FIM scores at admission and discharge after P4P was introduced, as well as the high return home rate, the possibility of an intentional increase in the number of patients whose conditions would be easily improved could not be excluded, although this could not be construed as the definite result of a patient screening practice. In addition, following the implementation of P4P, there was improvement in the process by which health care delivery was provided, but neither the FIM gain nor the return home rate was significantly higher after P4P was introduced. A similar result was obtained when the data were analyzed only by hospital cases for the entire period before and after P4P implementation.

Conclusion: No effect of P4P implementation on FIM at discharge or the return home rate was observed in this study. In the future, continuous monitoring of changes after implementation of P4P and detailed analysis of the possibility of deliberate patient screening will be needed. A more indepth examination of the indicators and incentive criteria currently in use is recommended.

Keywords: pay for performance, Functional Independence Measure, improvement, return home rate, Rehabilitation Patients DataBank of Japan

Introduction

The pay for performance (P4P) system has been implemented in many countries, including the US, UK, Australia, Taiwan, Korea, and Japan, in order to assess and manage the quality of health care systems properly, as well as to enhance their quality and effectiveness. Each of these countries has adopted P4P in combination with schemes such as fee-for-service, diagnosis-related groups, and the capitation system, to suit local health care systems.

There are growing concerns about health care quality in Japan. The introduction and implementation of P4P programs for recovery rehabilitation wards in 2008 was a significant advance in health care. In Japan, P4P in inpatient rehabilitation of stroke survivors began in 2008, with the primary objective of providing the services needed

Correspondence: Seungwon Jeong
Center for Well-being and Society,
5–22–35, Chiyoda, Nakaku, Nagoya
460–0012, Japan
Tel +81 52 242 3075
Fax +81 52 242 3076
Email k-jeong@n-fukushi.ac.jp

17

to improve functional recovery following stroke. Despite concerns about the effectiveness and quality of P4P expressed by some stakeholders in the program, three indicators were devised, ie, >60% of stroke survivors should be discharged into the community, >20% of hospitalized patients should be severe stroke cases, and >30% of patients in the P4P program should demonstrate an improvement in their performance of activities of daily living or their functional recovery at the time of hospital discharge.[1] If a hospital was able to satisfy all these requirements, it would be paid about 17,200 yen as an incentive. In 2010, 85% of recovery rehabilitation wards were receiving incentives.[2]

P4P is considered to be a tool that can enhance quality and improvement in health care. Nevertheless, there are no reports of formal evaluation of hospital P4P in the literature.[3] Given the different approach taken by each country regarding P4P, even with a significant increase in studies verifying the relationship between the implementation of P4P and the quality of health care systems,[4–7] the clear implication is that evaluation of the P4P adopted by each country should also take into account their unique situation.

Three years have now passed since the implementation of P4P in Japan, during which time several surveys have been performed to see whether or not patients were selected according to the criteria of the P4P scheme, but adequate verification of the effectiveness of P4P was not attempted. In this study, we evaluated the influence of P4P on the performance of hospitals by using data produced by the Rehabilitation Patients DataBank of Japan. Specifically, we investigated whether any improvement occurred in the Functional Independence Measure (FIM™)[8,9] at the time of discharge from hospital, the return home rate, and whether there was any intentional selection of patients expected to have greater recovery following treatment. We looked for any changes in terms of FIM improvement and patient return home rate which had taken place before and after implementation of P4P.

Materials and methods

The Rehabilitation Patients Databank of Japan[10] was developed to facilitate this evaluation, with financial support from the Ministry of Health, Labor, and Welfare of Japan. By May 2011, 37 hospitals had contributed structured data for a total of 9031 patients to the databank. It should be noted that not all the hospitals have a rehabilitation unit. In this study, we used only 3233 cases of recovery rehabilitation wards. Hospitals contracted to the databank collected patient data twice per year, first between January and February and then between July and August, from April 2006 to March 2010.

Patients discharged from hospital during the survey period were registered with the databank.

Data

To ensure our outcomes of interest were comparable across the different facilities, we divided the data into two groups, ie, before P4P (April 2006 to March 2008) and after P4P (April 2008 to March 2010). Data were included for patients aged 65–100 years who had a hospital stay of 77–180 days. Finally, the total numbers of patients and hospitals included were 903 and 26, respectively (before P4P [14 hospitals, n = 530] and after P4P [12 hospitals, n = 373]).

Analysis

Changes in patient characteristics and procedures undertaken before and after implementation of P4P were followed by analysis of changes in FIM score and return home rate. An assessment item for FIM was independently developed by the Japanese health care system, but the FIM used globally as a functional assessment measure of P4P[11,12] was used in the Rehabilitation Patients DataBank of Japan and also in this study.

Patient characteristics and processes before and after P4P implementation

To identify any changes in patient characteristics and procedures undertaken before and after P4P implementation, data on patient age, number of days from onset of stroke, and the modified Rankin Scale[13–15] score before the stroke, were analyzed by t-test. The modified Rankin Scale is commonly used to measure the degree of disability or dependence in the daily activities of people who have suffered a stroke, and it has become the most widely used clinical outcome measure for clinical stroke trials. The scale runs from 0–6 (0 = perfect health without symptoms, 6 = to death).[16]

The t-test was also used to analyze the FIM data. The FIM is an outcome measure of the severity of disability in an inpatient rehabilitation setting. It rates 18 activities of daily living on a seven-point scale ranging from fully dependent (1) to independent with no aids (7). The maximum total score is 126, indicating functional independence, and the lowest score 18, suggesting complete functional dependence. The items are grouped into two themes, ie, 13 motor items (eg, personal care, sphincter control, mobility, and locomotion) and five cognitive items (eg, communication and social cognition).[17] We measured the total FIM score at admission and at discharge, the FIM gain (FIM score at discharge – FIM score at admission), FIM efficiency ((FIM score at discharge – FIM score at admission)/length of stay in days), and amounts of

exercise prescribed by the physical therapist, occupational therapist, and/or speech/language/hearing therapist.

The Chi-square test was used to analyze data on the percentages of patients returning home (number of patients returning home/number of patients discharged), patients discharged without a caregiver, rehabilitation medicine specialists (ie, on the Japanese Association of Rehabilitation Medicine register) attending regular conferences, any interruption of more than 2 weeks of rehabilitation training, and diagnoses of cerebral infarction, cerebral hemorrhage, and subarachnoid hemorrhage.

Factors relating FIM improvement on discharge to P4P implementation

To investigate if there were any changes in FIM improvement-related factors before and after P4P implementation, three models were constructed, and multiple regression analysis was performed for input data. For model 1, the dependent variable was FIM score at discharge and the independent variables were patient age, motor and cognitive FIM scores at admission, number of days from onset of stroke, and modified Rankin Scale score before the stroke. In the next step, after considering factors such as interruptions to exercise of more than 2 weeks, patients with complications, and length of stay. Model 2 was constructed with the addition of the amount of exercise. Model 3 was constructed to investigate the impact of P4P, with hospitalization since 2008 (after P4P implementation) reflected in model 2 input as a dummy variable (after P4P implementation = 1).

Return home rates before and after P4P implementation

To assess the influence of P4P on factors related to return home, a logistic regression analysis was performed, with return home being the dependent variable (return home = 1) and the independent variables (from model 1) being patient age, modified Rankin Scale score before stroke, motor and cognitive FIM scores at admission, length of stay, amount of exercise, and availability of a caregiver. For admission since P4P introduction in 2008, an analysis was conducted to verify the influence of P4P (model 2) using a dummy variable (after P4P implementation = 1).

Results
Patient characteristics and processes before and after P4P implementation

There were statistically significant differences between the two groups for total FIM score at the time of admission,

length of stay in days, and percentages of patients discharged without a caregiver, rehabilitation medical specialist participation and regular attendance at conferences, interruption to exercise of more than 2 weeks, and amounts of exercise prescribed by a physiotherapist, occupational therapist, and/or speech/language/hearing therapist (Table 1). However, there were no statistically significant differences between the two groups for patient age, number of days from stroke onset, Modified Rankin Scale score before the stroke, percentage of patients returning home, total FIM score at discharge, FIM gain, FIM efficiency, or diagnosis (Table 1).

Factors relating FIM improvement on discharge to P4P implementation

Multiple regression analysis was used to investigate whether there were any changes in factors related to FIM improvement before and after P4P implementation. We found no statistically significant influence of P4P implementation (Table 2). In model 1, the younger the patient, the shorter the interval between stroke and admission, the lower the modified Rankin Scale score, the higher the motor and cognitive FIM scores at admission, with the FIM score at discharge being higher ($R^2 = 0.68$, $P < 0.001$). In model 2, the results indicated no change, even with input of more than 2 weeks and interruption to exercise, with complications regarded as compounding factors. In model 3, with P4P implementation as a dummy variable (P4P implementation = 1), a further analysis was performed, but no change could be attributed to P4P implementation ($R^2 = 0.66$, $P < 0.001$).

Return home rate before and after P4P implementation

A logistic regression analysis was performed to assess the influence of P4P implementation on factors related to return home (Table 3). In model 1, logistic regression analysis with a dependent variable of return home (return home = 1) showed that the younger the patient, the higher the motor and cognitive FIM scores at admission and, with the availability of a caregiver, the greater the likelihood of patients returning home. The Modified Rankin Scale score before the stroke, length of stay, and amount of exercise was not significantly different ($R^2 = 0.5$, $P < 0.001$). Using model 2, the analysis was performed with P4P implementation (P4P implementation = 1) added as a dummy variable ($R^2 = 0.54$, $P < 0.001$), but no change could be identified.

Table 1 Patient characteristics and processes before and after P4P implementation

Patient characteristics (t-test)	Before P4P (14 hospitals, n = 530)		After P4P (12 hospitals, n = 373)		P value
	Mean	SD	Mean	SD	
Patient age, years	76.2	6.6	76.2	7.1	NS
Days from stroke onset (n)	30.7	24.0	30.5	25.0	NS
Modified Rankin Scale score before stroke	3.6	1.0	3.5	1.3	NS
Total FIM score at admission	58.0	28.9	62.2	29.7	<0.05
Total FIM score at discharge	82.3	32.4	85.7	32.3	NS
FIM gain	24.2	20.1	22.6	18.9	NS
FIM efficiency	0.33	0.48	0.36	0.47	NS
Length of stay (days)	90.7	43.9	83.6	43.8	<0.05
Amount of exercise prescribed by PT (hours)	49.9	34.6	54.9	37.6	<0.05
Amount of exercise prescribed by OT (hours)	48.4	32.3	52.7	35.2	<0.05
Amount of exercise prescribed by ST (hours)	26.9	27.8	37.6	29.5	<0.05
Patient characteristics (Chi-square test)	**Before P4P (14 hospitals, n = 530)**		**After P4P (12 hospitals, n = 373)**		**P**
	%		%		
Patient return home rate	80.6		81.8		NS
Patients without a caregiver	22.0		29.6		<0.05
Rehabilitation medical specialist participation	26.4		48.3		<0.001
Regular conference attendance	37.6		84.1		<0.001
Interruption of exercise >2 weeks	4.9		1.1		<0.001
Diagnosis Cerebral infarction	76.3		73.9		NS
Cerebral hemorrhage	18.7		23.9		
Subarachnoid hemorrhage	5.0		2.3		

Abbreviations: P4P, pay for performance; PT, physiotherapist; OT, occupation therapist; ST, speech/language/hearing therapist; NS, not significant; FIM, Functional Independence Measure.

Discussion

Influence of P4P on changes in FIM score and return home rate

In this study, in order to identify any changes in hospital performance regarding rehabilitation as a result of P4P implementation in 2008, we used data from the Rehabilitation Patients Databank of Japan, in which indicators were applied to estimate outcomes which did not include structure or processes of quality of care. The extent to which changes in medical staff or individual hospital practices might have affected P4P implementation is not known.

In this study, some analyses were carried out to determine whether adoption of P4P has brought about any change in FIM scores for activities of daily living or the return home rate. Unfortunately, we could not demonstrate a relationship between P4P implementation and outcomes for rehabilitation medicine in Japan. However, our findings do suggest that P4P implementation has led to some changes in hospital performance. The length of hospital stay was shorter and FIM

scores slightly higher after P4P implementation. Moreover, there was an increase in the amount of exercise prescribed by physiotherapists, occupation therapists, and speech/language/hearing therapists, with more participation by medical specialists and increased conference attendance after P4P implementation, along with a decrease in the number of interruptions to exercise of more than 2 weeks, suggesting that treatment processes have improved. Any improvement in effectiveness is likely attributable to improved processes rather than patient-related factors. In the future, it will be necessary to verify whether these improvements have been prompted by P4P implementation, or by increased prescription of exercise.

Changes in patient characteristics after P4P implementation

According to research published in 2009,[18] there might have been some patient selection after P4P implementation. Considering that the FIM score at admission was found to

Table 2 Factors related to FIM improvement on discharge from hospital and influence of P4P implementation

Factor	Model 1		Model 2		Model 3	
	β		β		β	
Patient age	−0.09	***	−0.13	***	−0.13	***
Motor FIM at admission	0.42	***	0.40	***	0.40	***
Cognitive FIM at admission	0.45	***	0.44	***	0.44	***
Days from stroke onset (n)	−0.06	**	−0.07	**	−0.07	**
Rankin Scale score before stroke	−0.07	**	−0.06	*	−0.06	*
Any interruption to exercise >2 weeks (= 1)			−0.00	NS	−0.00	NS
No complications (= 1, having complications = 0)			0.02	NS	0.02	NS
Length of stay (days)			−0.08	NS	−0.07	NS
Amount of exercise (PT, OT, ST)			0.04	NS	0.03	NS
P4P implementation (= 1)					0.01	NS
R^2	0.68 (F = 327.7)***		0.66 (F = 110.8)***		0.66 (F = 99.6)***	

Notes: Dependent variable is total FIM score at discharge. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.
Abbreviations: P4P, pay for performance; PT, physiotherapist; OT, occupation therapist; ST, speech/language/hearing therapist; NS, not significant; FIM, Functional Independence Measure.

be higher in this study, it is possible that hospitals might have selected patients with a better chance of achieving improvement in FIM score. On the other hand, Japan's Central Social Insurance Medical Council reported in 2010 that it is impossible to determine clearly whether or not the selection of patients was intentional.[2]

In the light of the results of this study, to improve functional status and increase the return home following the implementation of P4P, one could consider the possibility of selecting patients with a higher FIM during hospitalization and increasing their exercise. In this study, the total FIM score at admission was significantly higher (4 points) after P4P than before its implementation. On the other hand, the proportion of patients without a caregiver increased in this time, leading to a decrease in the return home rate, so it is difficult to say whether there was any selection of patients for financial gain. This result is consistent with the findings of recent reviews of P4P and quality of the health care system,[19–21] with many researchers reporting that the suspicion of P4P implementation leading to gaming by hospitals is unfounded.

P4P implementation and improvement in health care systems

Implementation of P4P should promote evidence-based treatment guidelines that enhance the quality of the health care system, and criteria for measurement of these have been established elsewhere. However, in Japan, P4P was introduced as a social experiment[22] at the national level, as opposed to a model project in specified facilities and regions. This is in contrast with the US, where implementation of P4P took place in a piloted fashion, with studies investigating the relationship between implementation of P4P and quality of health care have been conducted for the past two decades.

The studies reported by the US and other countries until 2009 show that P4P implementation has not necessarily

Table 3 Change in return home rate before and after implementation of P4P

	Model 1				Model 2			
	B	Wald	Odds ratio		B	Wald	Odds ratio	
Patient age	−0.03	5.46	0.97	*	−0.01	0.77	0.99	NS
Modified Rankin Scale score before stroke	−0.17	1.51	0.85		−0.37	4.34	0.69	*
Motor FIM at admission	0.07	32.61	1.06	***	0.06	24.50	1.07	***
Cognitive FIM at admission	0.07	15.20	1.07	***	0.10	16.39	1.10	***
Length of stay in days	0.00	0.01	1.00	NS	0.00	0.16	1.00	NS
Amount of exercise (PT, OT, ST)	0.05	0.53	1.06	NS	−0.03	0.11	0.97	NS
Along with the availability of a caregiver (= 1, none of caregivers = 0)	1.30	14.58	3.67	***	1.23	10.05	3.43	**
P4P implementation (= 1)					−0.39	1.28	0.68	NS
Nagelkerke R^2		0.51***				0.54***		

Notes: Dependent variable is return home (= 1, otherwise = 0). $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. Dependent variable is total FIM score at discharge $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.
Abbreviations: P4P, pay for performance; PT, physiotherapist; OT, occupation therapist; ST, speech/language/hearing therapist; NS, not significant; FIM, Functional Independence Measure.

resulted in improved quality of health care.[3,4,23,24] Furthermore, because P4P is being implemented in health care systems that differ by country and domain, it is necessary to evaluate their effect individually. In the UK, reports claiming greater equity in the health care system now exist.[25–27]

P4P has a short history, and more time may be needed before its effects become visible. It is possible that choosing an appropriate program to use is important, rather than introducing P4P in a generic way. Accordingly, in Japan, more specific verification of quality improvement regarding type of medical institution, patient, and health care system is required, and, at the same time, there is a need for continued investigation to identify appropriate indicators for evaluation of P4P.

In this study, we could not determine the effect of P4P implementation on FIM score at discharge or on the return home rate. Our sample size and follow-up period may have been inadequate for verifying the impact of P4P. However, a reliable and relevant database does not exist in Japan. Under these circumstances, identification of any change in patient characteristics after the implementation of P4P by using data from various hospitals can be seen as meaningful. In the present study, an overall analysis was conducted in all hospital facilities before and after implementation of P4P, but a more thorough analysis is required to assess the performance of each hospital receiving P4P incentives.

Conclusion

This study analyzed data from the Rehabilitation Patients Databank of Japan, which holds patient records from various hospitals, in order to identify any change in patient care after P4P was introduced and to clarify the effect of P4P implementation on FIM score and the return home rate. Given the high FIM scores at admission and at discharge after P4P was introduced, as well as the high return home rate, the possibility of treatment being targeted to patients whose conditions can be easily improved cannot be excluded. However, at the same time, this cannot be construed as a definite result of a patient screening practice, and it is possible that P4P implementation has achieved positive changes in hospital performance. In this study, different samples of hospitals were used for analyses before and after P4P implementation – 12 and 14 hospitals, respectively. However, a similar result was previously obtained for four hospitals, when an intensive analysis was conducted using the data during the entire period both before and after P4P implementation. In the future, ongoing monitoring of changes after implementation of P4P and detailed checks for possible patient screening will

be needed. A more indepth examination of the indicators and incentive criteria in use will also be necessary.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Ministry of Health, Labor and Welfare of Japan. Available from: http://www.mhlw.go.jp/english/index.html/. Accessed September 16, 2011. [Japanese].
2. Japan's Central Social Insurance Medical Council. An investigative report on the effect of 'quality assessment' introduced in relation to the rehabilitation ward hospitalization fee. Japan's Central Social Insurance Medical Council. 2010. [Japanese].
3. Jeong S, Kondo K, Inoue Y. Pay for Performance in USA and suggestion to Japan. *Shakai Hoken Jyunpo*. 2009;2397:20–25. [Japanese].
4. Christianson JB, Leatherman S, Sutherland K. Lessons from evaluations of purchaser pay-for-performance programs: a review of the evidence. *Med Care Res Rev*. 2008;65(6 Suppl):5S–35S.
5. Jha AK, Orav EJ, Dobson A, Book RA, Epstein AM. Measuring efficiency: the association of hospital costs and quality of care. *Health Aff (Millwood)*. 2009;28(3):897–906.
6. Lindenauer PK, Remus D, Roman S, et al. Public reporting and pay for performance in hospital quality improvement. *N Engl J Med*. 2007;356(5):486–496.
7. Rosenthal MB. Beyond pay for performance—emerging models of provider-payment reform. *N Engl J Med*. 2008;359(12):1197–1200.
8. Granger CV, Deutsch A, Linn RT. Rasch analysis of the Functional Independence Measure (FIM) mastery test. *Arch Phys Med Rehabil*. 1998;79(1):52–57.
9. Keith RA, Granger CV, Hamilton BB, Sherwin FS. The functional independence measure: a new tool for rehabilitation. *Adv Clin Rehabil*. 1987;1:6–18.
10. Jeong S, Kondo K, Shiraishi N, Inoue Y. An evaluation of the quality of post-stroke rehabilitation in Japan. *Clinical Audit*. 2010;2:59–66.
11. Ottenbacher KJ, Hsu Y, Granger CV, Fiedler RC. The reliability of the functional independence measure: a quantitative review. *Arch Phys Med Rehabil*. 1996;77(12):1226–1232.
12. Hamilton BB, Laughlin JA, Fiedler RC, Granger CV. Interrater reliability of the 7-level functional independence measure (FIM). *Scand J Rehabil Med*. 1994;26(3):115–119.
13. Van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJ, van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke*. 1988;19(5):604–607.
14. Banks JL, Marotta CA. Outcomes validity and reliability of the modified Rankin scale: implications for stroke clinical trials: a literature review and synthesis. *Stroke*. 2007;38(3):1091–1096.
15. Quinn TJ, Dawson J, Walters MR, Lees KR. Reliability of the modified Rankin Scale: a systematic review. *Stroke*. 2009;40(10):3393–3395.

16. Rankin J. Cerebral vascular accidents in patients over the age of 60. II. Prognosis. *Scott Med J*. 1957;2(5):200–215.

17. Uniform Data System for Medical Rehabilitation. About the FIM system. Available from: http://www.udsmr.org/WebModules/FIM/Fim_About.aspx. Accessed September 16, 2011.

18. Japanese Association of Rehabilitation Medicine. Report for the relation to the rehabilitation ward hospitalization fee. *The Japanese Journal of Rehabilitation Medicine*. 2009;467. [Japanese].

19. Van Herck P, De Smedt D, Annemans L, Remmen R, Rosenthal MB, Sermeus W. Systematic review: Effects, design choices, and context of pay-for-performance in health care. *BMC Health Serv Res*. 2010;10:247.

20. Doran T, Fullwood C, Gravelle H. Pay-for-performance programs in family practices in the United Kingdom. *N Engl J Med*. 2006; 355(4):375–384.

21. Ryan AM. Effects of the premier hospital quality incentive demonstration on Medicare patient mortality and cost. *Health Serv Res*. 2009;44(3):821–842.

22. Niki R. Health care reform and rehabilitation medicine in the future. *Community Rehabilitation*. 2008;44:234–242. [Japanese].

23. Mehrotra A, Damberg CL, Sorbero ME, Teleki SS. Pay for performance in the hospital setting: what is the state of the evidence? *Am J Med Qual*. 2009;24(1):19–28.

24. Greene SE, Nash DB. Pay for performance: an overview of the literature. *Am J Med Qual*. 2009;24(2):140–163.

25. Ashworth M, Seed P, Armstrong D, Durbaba S, Jones R. The relationship between social deprivation and the quality of primary care: a national survey using indicators from the UK Quality and Outcomes Framework. *Br J Gen Pract*. 2007;57(539):441–448.

26. Crawley D, Ng A, Mainous AG, 3rd, Majeed A, Millett C. Impact of pay for performance on quality of chronic disease management by social class group in England. *J R Soc Med*. 2009;102(3):103–107.

27. Karve AM, Ou FS, Lytle BL, Peterson ED. Potential unintended financial consequences of pay-for-performance on the quality of care for minority patients. *Am Heart J*. 2008;155(3):571–576.