HEALTH CARE MANAGEMENT

Performance-based payment systems for general practitioners and specialists in selected countries: a comparative study

MOHAMMAD EBRAHIM EGHBALI[1], HAMID POURASGHARI[1,2], HASAN ABOLGHASEM GORJI[1,3], MARIANO MARTINI[4], JALAL ARABLOO[3], MASOUD BEHZADIFAR[5], AIDIN ARYANKHESAL[11]
Department of Health Services Management, School of Health Management and Information Sciences, Iran University of Medical Sciences, Tehran, Iran; [2] Hospital Management Research Center, Health Management Research Institute, Iran University of Medical Sciences, Tehran, Iran; [3] Health Management and Economics Research Center, [4] Health Management Research Institute, Iran University of Medical Sciences, Tehran, Iran; Department of Heatlh Sciences, University of Genoa, Genoa, Italy; [5] Social Determinants of Health Research Center, Lorestan University of Medical Sciences, Khorramabad, Iran

Keywords

Pay for performance • Quality-based payment • Value-based payment • Performance evaluation indicators • Health policy

Summary

Background. *Due to the growing increase in the needs of health systems in the field of financial and human resources management, performance-based payment has been the subject of attention by health and welfare policymakers. This study aimed to compare the components of performance-based payment in selected countries.*

Methods. *This comparative study was conducted in 2021. The selection of countries was based on three measures: the type of health insurance system, the development of the performance-based payment system, and the state of economic development of the countries. The findings were organized using comparative analysis tables. The general framework of performance-based payment systems, including goals, activities and actions, people involved in the program, and the way of encouraging and punishing, was used for analysis.*

Results. *The findings of the study showed that in most of the programs, aspect of clinical quality has the highest weight. Other dimensions include patient experience and satisfaction, physician financial performance, and patients' access to services. In most programs, various risk adjustment methods such as exception reporting, combined payments, payment according to demographic characteristics, were used to reduce provider risk, and clinical service providers were actively involved in the program design progressive.*

Conclusions. *Despite the widespread use of performance-based payment programs in most countries, these programs face limitations and shortcomings. By linking incentives to individual, team, and organizational performance, a performance-based payment program can improve teamwork, and create integrated health care.*

Introduction

Health organizations have a special position in the society due to their duties in the field of prevention, care, treatment and rehabilitation, and any weak performance in their management causes a delay in timely treatment and disease progression or death [2, 1]. Given that in hospitals, human resources play a main role as the core of the organization, inefficient payment systems and insufficient salary or wage has led to some problems, such as absence at work, employee dissatisfaction, conflicts between employees, quitting the job, strike, and complaints [3, 4]. Having a motivated and competent workforce and a performance evaluation system increases the effectiveness and efficiency of services in health care delivery [5, 6]. Performance appraisal seeks to find the most accurate and cost-effective methods for measuring job performance and job satisfaction [7, 8]. Considering that financial incentives are one of the most important factors affecting individual and organizational behavior in the health sector and have many effects on the health system and the quality and quantity of services, health sector managers should consider the powerful effects of motivation on employees' behavior in designing a payment system [9, 10]. According to Steven's study, giving employees 10% more bonuses can be motivated them [11]. Pay for Performance (P4P) is a payment model that attempts to reward measured dimensions of performance and incentivize health service providers to achieve predetermined goals by financial incentives [12, 13]. P4P was created in order to improve quality and efficiency and reduce additional costs by providing financial incentives to payers and service providers to establish a relationship between economic incentives and the quality of their performance [14]. P4P is different from other traditional payment methods. Traditional methods relate income to workload and do not consider paying for quality, while P4P explicitly addresses efficiency and effectiveness [15]. The most

 https://doi.org/10.15167/2421-4248/jpmh2025.66.1.3511

important advantage of P4P is that they significantly improve the organization's performance without putting additional financial pressure on it [16]. P4P programs intend to improve the system performance and accountability by motivating health workers and increasing their independence in employing available resources [17]. Based on the principal-agent theory, P4P is expected to align payers' and providers' motivations [18]. Motivation, defined as an individual's willingness to try and act towards organizational goals [19]. Other benefits of P4P system include attracting potential job seekers, retaining employees, motivating employees, paying according to legal regulations, curbing organizational costs, and simplifying strategic goal [20-22].The P4P in different countries has different results. For example, implementing this system in the UK, which has one of the most comprehensive P4P programs at the primary care level in the world, has improved uterine screening tests and immunization, as well as improving the quality of the services provided before the introduction of the program [23]. In Turkey, P4P has also led to an increase in the productivity of doctors and a decrease in the number of patients per physician [24]. In general, implementing this system in Turkey has been satisfactory and has led to an increase in the quality of health services [25]. Contrary to the successful experiences in the implementation of this system, in some countries, the implementation of this system has faced problems, the main reasons for this are the costly and time-consuming implementation of the program, the resistance of healthcare workers to change, and the lack of Adequate infrastructure prior to implementation [26, 27]. For example, in Canada, due to the lack of an accurate performance evaluation system in many healthcare organizations, these organizations were not successful in applying performance-based payment methods [28].P4P has been widely implemented in developing countries over the past decade and has shown favorable results despite the existence of little evidence. Evidence has shown that P4P has led to improve chronic disease care [29], reduced hospitalization and mortality [30], improve documented care processes [31] and cost savings [32]. Insurance organizations, as service buyers and custodians of the payment lever in the health system, are obliged to move in the direction of connecting payments with performance quality. If an important reward such as payment cannot be linked with results, it will lead to a decrease in motivation and a drop in performance [33]. On the other hand, deterrents for violators should be designed large enough to encourage providers to invest in order to achieve performance goals [34].Without financial penalties and punitive options (for non-compliant and delinquent providers), P4P will only increase health costs [35]. Considering that P4P is a step towards achieving quality [36], this study was conducted with the aim of comparing the components of the P4P system in selected countries. This study has tried to identify the organization of P4P systems in selected countries to help planners and policy makers in designing a sustainable and effective P4P system.

Methods

Search strategy and selection process

This comparative study was conducted in 2021. In this study, a five-step protocol including determining the countries to be studied, searching for relevant documents, selecting documents, extracting data, and reporting, was used to conduct this comparative review. Searching for electronic resources based on the keywords of "Performance-based payment", "Quality-based payment", "Outcome-based payment", "Value-based payment", "Performance evaluation indicators" and "Quality evaluation indicators" and selected countries was conducted. These keywords were combined with Scopus and PubMed databases, Google Scholar, Google search engine, as well as the websites of the Ministry of Health, the World Bank, OECD, and WHO. The inclusion measures for the study included all articles and documents published between 2000 and 2021, relating to performance - based pay in selected countries.

The measures for entering the countries into the study included the type of health insurance system, the extent of using the performance-based payment system, and the economic development status of the countries. Also, having successful experiences and policies in implementing the performance-based payment system and having valid evidence in the selection of countries were considered. Finally, eight countries, including England (two programs), Taiwan, the United States of America, Canada, Germany, Turkey, France, and Iran, were selected to compare the components of the P4P system.

Comparative review of programs

The selected programs have been examined based on three main questions in accordance with the purpose of the research. In the first question (what is encouraged?), programs were evaluated in terms of performance dimensions and measures, measurement of indicators, provider participation, data collection, and methods used to adjust the risk of providers. In the second question (who is encouraged?), the programs were evaluated in terms of individuals or groups, the characteristics of the providers and the type of their participation (voluntary or non-voluntary). In the third question, in addition to evaluating the basic payments, the type (positive and negative), the amount, the period and the method of calculating the incentive and punitive payments were reviewed.

After reviewing the literature, the P4P system variables were identified, and the data were gathered using a researcher-made checklist based on the P4P framework. The researcher-made checklist contained all the information related to the objectives of the study. The extracted data were classified according to the components of the analysis and were then organized into comparative tables. Three comparative table was completed for the eight selected countries (nine programs). The comparative tables included components such as performance indicators, providers and financing

methods in selected countries. For this purpose, similarities and differences between the countries were compared based on the information extracted from the comparative tables. Framework analysis was used to analyze the data, and the data analysis was performed using comparative analysis tables, which compare the components of P4P system.

Results

In this research, nine programs were selected among the P4P programs in the world, which are: Quality and Outcomes Framework (QOF) and Advancing Quality (AQ) in England, National Health Insurance_Pay for performance (NHI-P4P) in Taiwan, Integrated healthcare association-physician incentive program (IHA-PIP) in the California of US state, Physician Integrated Network (PIN) in Canada, Ergebnis Orientierte Vergütung (ERGOV) in Germany, Family medicine performance based contracting scheme (FM PBC) in Turkey, Rémunération sur Objectifs de Santé Publique (ROSP) in France, and Payment guidelines based on the performance of medical practitioners and faculty members working in university-affiliated hospitals (P&FM-P4P) in Iran.

What is encouraged?

Six programs were initiated by a public purchaser and three by private insurers responsible for managing the care for their enrollees. In the programs targeting multiple dimensions, clinical quality has the highest weight and the most scales were included. Other dimensions include financial performance, patient satisfaction/experience, capacity and access (for example, structural measures referring to administrative and organizational aspects of performance such as receiving training/providing and record keeping). Clinical aspects typically relate to chronic and preventive care, However attention to acute care is considered in some programs such as AQ, FM-PBC, NHI-P4P, QOF and P&FM-P4P. Four programs adopted a set of measures including at least 30 measures pertaining to clinical quality and patient satisfaction/ experience or access.

Various other risk-mitigation methods are used across the programs. Risk adjustment is used in AQ, ERGOV and ROSP for financial purposes. Especially in AQ and ERGOV, risk adjustment appears to be relatively complex, controlling for sociodemographic and morbidity-based risk factors. Not all programs that include outcomes apply risk adjustment. Various other risk mitigation methods are used throughout the programs. For example, in IHA-PIP, performance targets are differentiated based on how current performance is influenced by case mix and population characteristics. In general, although the documents provide limited information on the use of risk-mitigating measures, the results raise doubts about whether differences in risk are sufficiently equalized, especially in PIN and NHI-P4P. In P&FM-P4P, hospitals with low financial performance, add 5% of the hospital's income (Institute share) or the institution's aid from 5% of other hospitals, to the physicians' contribution limit. In most programs, providers actively participate in program design and implementation. The participation of the provider is considered as a critical success factor(37), and is being realized in various ways, including delegating authority to providers (QOF, FM PBC), consensus meetings (AQ, IHA-PIP) and using feedback from providers (QOF, NHI-P4P, FM PBC). Table I evaluate the various performance programs and what is encouraged.

Who are encouraged or punished?

In most programs, payments are mostly provided at the group level. "Targeted groups" vary in structure and size, ranging from hospitals (AQ, FM PBC) to large multispecialty organizations (IHA-PIP) to primary care efforts (QOF, ROSP). In ROSP, payment is provided to the primary care practice for measures for which this does not seem to hold. For example, GPs receive more money for each Pap smear, but if a specified percentage of patients are screened, the physicians receives a fixed amount per patient [67]. In the NHI-P4P, payment is provided to hospitals for cancer and diabetes, but directly to physicians for TB and asthma. However, for many measures included in these program, sample sizes may not be sufficient to generate reliable profiles, particularly for outcomes and resource use [68].This also seems relevant for PIN and QOF, as many physicians still work in small group-practices. For several programs, data state that measures are only included if they are adequately under providers' control and/or if sample size is adequate. However, it is not clear when this is the case. Some programs (*e.g.,* AQ, P&FM-P4P, ROSP) aggregate individual measures into composites, which can increase reliability [69, 70]. In ROSP this resulted in fair reliability, despite that many physicians were duo or solo practices. Although it is difficult to draw conclusions, there are concerns about whether providers can be sufficiently discriminated from each other, and thus whether payment allocation occurs sufficiently. In P&FM-P4P, the performance of each doctor is calculated and paid individually. Participation in this program is non-voluntary and the calculation and payment of the performance of all doctors who have contracts with government health insurances is done under the terms of this plan [71].

In five programs (AQ, FM PBC, ROSP, QOF, P&FM-P4P) the participation rate is virtually 100 percent. In PIN, IHA-PIP and NHI-P4P participation is more than 50%. Low participation in ERGOV may be problematic. In ERGOV, clinics participating in the scheme are known as preferred providers. This may be a strong incentive for clinics to participate, especially if receiving care from unlicensed providers requires large out-of-pocket payments. But participation may still be unattractive because it involves a considerable administrative burden while financial consequences are highly uncertain. To achieve meaningful differences, participation must be increase [54]. Also in NHI-P4P for breast cancer, low

Tab. I. What are the details of the various P4P programs and what is encouraged?

Name of the program/ country/year of commence /number of measures	Performance dimensions (weight)	Performance measures	Development and evaluation method	Methods used to mitigate providers' risk
QOF United Kingdom 2004 134 measures (23, 38-40)	• Clinical (69.6%) • Organizational (16.7%) • Patient experience (9.3%) • Additional services (4.4%)	• Clinical: 86 measures, 20 areas (chronic, acute, prevention, psychological) • Organizational: 36 measures • Patient experience: 3 measures • Additional services: 9 measures	• Evaluation, review and development of Measures by professional organizations • Selection/weights based on negotiations between the government and the British Medical Association • Data collection: uniform electronic medical record managed by General Practitioners, extracted to national database	• Annual inspections by primary care centers, big penalties for fraud • Rejection of some patients by general practitioners. Exception reporting
AQ United Kingdom 2008 At least 30 measures (30, 39, 41)	• Clinical (60%) • Patient experience (20%) • Patient-reported outcome measures (PROMs) (20%)	• Clinical: 27 processes, 3 final outcomes; divided over 5 acute care areas • PROMs: quality of life before and after medical services	• Measures developed within the framework of CMS/Premier Hospital Quality Incentive Demonstration in the US • Self-collected data; goals for completeness/accuracy; centralized support • Endorsed by royal colleges and clinicians.	• Risk-adjustment: survival index for acute myocardial infarction, PROMs • Composite score for each therapeutic and clinical domain
NHI-P4P Taiwan 2004 At least 22 measures (42-46)	• Clinical (100%) Diabetes mellitus, breast/cervical cancer, asthma, tuberculosis • In order to add clinical outcomes, hepatitis B/ C, schizophrenia, and hypertension	• Diabetes mellitus: 2 structures, several processes, 2 intermediate outcomes • BC: 4 structures, processes, 2 outcomes • Asthma: 2 structures, several processes • TB: 4 structures • CC: 2 processes, several processes, 1 final outcome	• Data self-reported by providers and automatically entered into the database • Measures are selected based on disease burden • Intention to increase participation of providers in program development and measure selection	• Requires sample size. • Providers decide which patients to admit. Government increases the number of patients physicians have to admit (in 2010 for DM: 33% of population, ≥ 55 patients)
IHA-PIP America-California 2003 Four main dimensions with at least 25 measures (47-50)	• Clinical quality (50 percent) • Patient experience (20 percent) • Effective use of health information technology (30%) • Appropriate use of resources (without percentage)	• Clinical quality: preventive care, childbirth, cardiovascular, diabetes, musculoskeletal and respiratory • Patient experience: communication with the doctor, coordination of care, timely access, privacy, respect and quality of service delivery, • Appropriate use of resources: utilization of outpatient and inpatient services, referral to the emergency room and prescription of generic drugs	• Measures developed/ negotiated at the national level; Doctors are consulted • The provider payment formula is developed through negotiation with the government and physicians • Separate storage of patient records	75% of payments are determined according to the age and gender of the patient
PIN Canada, Manitoba 2004 37-40 measures (51-53)	• Clinical (100%) • Intention to add ongoing care, access, mental health, and coordination	• 24 processes and 6 areas for chronic care • 14 processes for preventive care • 2 additional processes for depression care	• Selection: expert opinion, consensus meetings • Data system populated via clinics' electronic records	• Measures included if "specific" to clinics and data are valid and Reliable • Checks with registry • Measures adjusted based on feedback

▶

Tab. I *(follows).* What are the details of the various P4P programs and what is encouraged?

Name of the program/ country/year of commence /number of measures	Performance dimensions (weight)	Performance measures	Development and evaluation method	Methods used to mitigate providers' risk
ERGOV Germany, regional 2001 20 item tool (54, 55)	• Patients' ability to perform daily life activities • Quality of outcome of rehabilitation care for stroke patients	• Self-care (7 items), mobility (4), communication skills (4), cognitive activity (5) • 6 types of help in each case	• A quality assessment tool combining items from widely used measurement tools with good psychometric properties • Approved by clinics	• Data is reported online. Checks using self-assessment • Rejection of the patient • ≥100 patients
FM PBC Turkey 2006 35 measures (56-59)	• Clinical (55%) • Services (25%) • Management evaluation (20%)	35 performance measures including: 19 clinical cases Services: 9 items and management evaluation: 7 items	Selection based on experience from other countries, clinical communication and data collection through electronic health records is possible	Creating follow-up or reminder lists for family physician staff as a decision support system
ROSP France 2012 29 measures (60-62)	• Clinical (58%) • Appropriate and efficient versions (24%) • Quality management (17%)	• Chronic disease management and follow-up (9 measures) • Prevention (8 measures) • Appropriate and efficient versions (7 measures) • Office organization and quality of care (5 measures)	• Selection based on disease burden, consensus and available evidence • Development by clinical networks and verification of information by quality measurement. • Information was collected by clinical units	• Actions are only taken if they are sufficiently within the providers' control
P&FM-P4P Iran 2013 27 measures (63-66)	• Clinical measures (46.4%) • Patient satisfaction (32.2%) • Organizational measures (21.4%)	• 13 performance measures for specialist doctors • 9 measures to measure the performance of general practitioners • 6 measures to measure patient satisfaction	• Choosing the weights based on the decisions of Council for planning and supervising the distribution of special income • Data collection: electronic medical records extracted in the database	• Measurements are explicitly selected based on sufficient sample size in each treatment center. • The fixed part of the fee of the program to support the longevity of doctors in deprived areas is not included in the calculation of the professional component

participation appears to be a result of the additional financial risk that participation involves and the fact that hospitals experience survival rates, which determine whether or not they receive a reward, to be largely beyond their control [45].Table II provides more detail on "who is encouraged or punished".

How is encouragement and punishment done?

In AQ, there were no penalties for poor performers, but hospitals that failed to meet targets for data accuracy and completeness received a penalty or were removed from the program. (The current version of AQ involves withholding of payments rather than rewards) [41]. Compared with ERGOV, the financial risk of participation was lower. Also, there was less uncertainty because payments were fixed. Due to the size of the payment, there is much variation across programs. In AQ, in addition to payments for patient-reported outcomes, hospitals could receive a 4 percent add-on to the national tariff for the associated activity. In FM PBC, bonus potential is 5 to 10 percent of the average DRG price. In NHI-P4P, payments per patient are often maximized per year. For cervical cancer, fees may be increased by up to 55 percent. For patients with breast cancer, qualified hospitals receive a bundled payment, which is higher than the regular payments. Hospitals also meeting targets for disease-free survival are qualify for a bonus of up to 8 percent of the bundled payment [44]. In ERGOV, clinics are judged based on their performance relative to the mean. In addition, only three clinics in each of the five clinics that best achieve their own target receive a bonus proportional to the degree of target attainment. Five programs use three or more targets or a sliding scale [55]. FM PBC typically uses five targets per measure with a large difference between levels. A similar approach is used in ROSP, which may well have contributed to the finding that improvements in incentivized measures were typically largest among GPs with medium or low baseline performance [72]. QOF and P&FM-P4P use a sliding scale. Providers in QOF earn

Tab. II. Who are encouraged or punished in the performance-based payment program of selected countries?

Program name	Characteristics of providers	Individual or group?	Type of participation
QOF	• 8,600 primary care practices (almost 100%) • On average 5,500 patients, 3.6 physicians • Gatekeeping and patient enrolment are mandatory	Mostly group, but individual activity is also possible (6% of total in 2008)	Voluntarily
AQ	• All 24 hospitals in the Northwest region of England that provide emergency care • Hospitals can be public or private	• Group • Payments allocated to clinical teams to invest in patient care	Voluntary
NHI-P4P	• Diabetes: Hospitals. (Physicians' participation: 47 percent) • Asthma: pediatricians, internists and general practitioners • Breast/cervical cancer: hospitals • Tuberculosis: hospitals (43% participated in 2006),	• Diabetes, breast/cervical cancer: in groups • Asthma: individual • Tuberculosis: both	Voluntary
IHA-PIP	All 1500 health care clinics	Medical groups	Voluntary
PIN	• Phase 1: typically 15 to 30 physicians (mostly general practitioners, but also specialized physicians and other practitioners) • Currently (phase 2): 14 primary care groups • Measures to participate: electronic medical record, ≥5 GPs, 6,600 patients, access for other general practitioners	• Group, but payment often divided over participating physicians • Primary care groups receive funding from member clinics • Free budget allocation by clinics	Voluntary
ERGOV	• 13 rehabilitation clinics (pilot). • Project with the financial support of clinics • Clinics assess patients at admission and at discharge	• Group	Voluntary
FM PBC	Family doctor units	• Family doctor team	Voluntary
ROSP	72 general practitioners (1.7% of all doctors)	• Individual • Group • Organizational	Voluntary
P&FM-P4P	General and specialist doctors, medical fellowships, dentists, faculty members and assistants included in this directive in the hospital and special clinic affiliated to the university/faculties of medical sciences and health services of the Ministry of Health, Treatment and Medical Education	• Individual • Group	Not voluntary

more bonus for a larger percentage improvement from baseline to the goal. In QOF, each measure has upper and lower targets delineating the scale. Performance improvement were most pronounced for GPs with low scores at baseline, which could have been a result of the sliding scale on which practices are scored [73]. In P&FM-P4P, each performance measure has low and high targets that define the scale [74].

. Providers receive additional revenue for increasing a higher percentage of baseline to service delivery goals. In NHI-P4P provides piece rates for process quality. For example, for breast cancer, hospitals are rewarded for each patient completing recommended cure. This may well have contributed to observed improvements in process quality and the result [45].

Although the performance goals in most programs are different, they follow one or more common goals, including "improving the quality of care", "increasing patient satisfaction" and "improving processes". For example, the main goal in P&FM-P4P, IHA-PIP, QOF and ERGOV programs is to improve the quality of service coverage and improve the individual and team performance of providers. In the NHI-P4P program, ROSP, PIN, improving preventive care is a major goal. Table III provides more details on "how to encourage and punish".

Discussion

The present study provides an international overview of P4P initiatives in health care. The nine identified programs have similar design in several respects. All encourage clinical quality and most of them only use positive incentives, actively involve providers in design, and based on performance, they pay monthly, quarterly or annually. However, there is also considerable heterogeneity regarding the breadth of measure sets, use of risk-mitigating measures, number and type of targets and payment size. there seems to be ample room In most programs to increase incentives for desired behavior and to mitigate incentives for undesired behavior. In particular, shortcomings pertain to number and type of included performance measures, risk adjustment of outcomes and resource use, payment frequency, reliability of measurements, and number of targets.

Different P4P programs in the world consider different dimensions of performance and include different indicators in the program. Experts believe that the indicators of the P4P program should be valid and consider the areas of process and outcome together with each other [75].

Modification seems relevant mainly for ERGOV, FM PBC, and NHI-P4P, but also for other programs there is

Tab. III. How to encourage and punish in the performance-based payment program in selected countries?

Program name	Performance goals	Encouragement or punishment	Incentive size	Basic payments	Payment calculation	Payment frequency
QOF	For each measure: sliding scale within absolute targets (typically 45% and 90%)	Just encouragement	Up to 30% of practice income	Risk-adjusted capitation	• Scores converted into points and then summed (up to 1000 points) • Fixed amount per point (£120), fixed no. Points per measurement	Annually
AQ	• 1st year, clinical: relative, +4% or +2% for reaching top or 2nd quartile of achievement • 2nd year, clinical: relative; attainment (1 target), most improved (1 target), top performance (2 targets)	• Encouragement • Penalty for inaccurate/ incomplete data	• Clinical: 3-5% additional to the tariff (bonuses totaling £3.3M in the first year and £1.7M in the second year) • Outcomes and patient reported experiences: both £1M/year • Max. between £260K-702K/year depending on the size of the hospital	National tariffs for clinical conditions	Hospitals ranked on composite score per clinical area, measures are equally weighted	Annually, often with a delay of 2 or 3 months
NHI-P4P	• Positive scores on structures • For Asthma, DM, BC, TB: enlarged fees for processes • Cervical cancer: number of and monthly growth in Pap smears • Diabetes: relative target for outcome measures • Breast cancer: absolute target for disease-free survival • Tuberculosis: cure rates	•Just encouragement • Financing not from global budgets	• Asthma: NT$1,3K/ patient/year • Cervical cancer: 15-50% add-on to current fees • Diabetes: NT$1,9K/ patient/ year for process measures • Breast cancer: 2, 3, 4, 6 and 8% add-on to bundled payment for 1, 2, 3, 4, and 5 year disease-free survival; on average NT$127K • Tuberculosis: hospitals NT$13K/ case; physicians NT$1,6K/ case; case manager NT$6K (first 6 months $3K, then $550/month)	• Fee-for-service under global budgets. 52 procedures reimbursed through fixed case payments	• Diabetes: outcomes: top 25% on composite • Breast cancer: stage-specific targets for survival, descending targets and ascending payment over the five survival-years • Cervical cancer: 15-50% add-on to current fees based on size of improvement •Tuberculosis: payment varies by 5 treatment stages and is larger if cured earlier	Monthly to annually
IHA-PIP	Medical organizations must operate above 50% of the three main areas of IHA indicators: clinical quality, patient experience and total cost of care	Just encouragement	Depends a lot on size of care / number of patients seen	Each insurance has its own budget and determines the method of calculating bonus amounts to its medical groups	The scores of each medical group in achieving and improving the index are added together, and then the total is multiplied by the corresponding weight.	Annually
PIN	• Phase 1: typically 4 absolute targets/ measure, large range (*e.g.*, 40, 60, 80, and 90%) • Phase 1: only 2 areas of performance measures	Just encouragement	Phase 1: data management: C$5K/ clinic, max. C$5K/ GP, C$360K in total. Group: C$40K/ clinic, max. C$5K/GP, C$370K in total	Fee-for-service	Phase 1: increasing payment per measure if higher goal is reached (*e.g.*, 50, 65, 80, 95, 100% of maximal payment per measure)	After the demonstration period (last quarter of 2008 for phase 1)

▶

room for improvement, notably regarding measure sets, payment size and risk-mitigating methods [76, 77]. For some aspects design seems sufficient in most programs. These include provider involvement in design (seven programs), voluntary participation (eight programs), and type of targets (absolute targets or piece-rates in eight programs). QOF and AQ seem to have been designed particularly well. The effectiveness of QOF has been evaluated in several studies [40, 78]. The most comprehensive program is QOF, which includes more

Tab. III (*follows*). How to encourage and punish in the performance-based payment program in selected countries?

Program name	Performance goals	Encouragement or punishment	Incentive size	Basic payments	Payment calculation	Payment frequency
ERGOV	Quality tournament in which rewards and maluses are determined by relative differences (deviation from mean)	• Both encouragement and punishment • Until now, payments have been virtual	• Depends on how clinic performs relative to other clinics and on size of coefficient (can be adjusted so that clinics do not go bankrupt) • Neutral funding: maluses for low-performers used to finance bonuses for high-performers	Typically a daily or flat payment per patient	• 100 points for each patient • Value at discharge subtracted from predicted value, calculated using admission scores from all patients. The remaining amount per clinic averaged and multiplied by no. of patients to calculate bonus	Quarterly
FM PBC	The contract framework includes two performance levers: • Salary deduction • Warning points	Often punishment	The service credit is calculated in the table that can be adapted to the income of the people and can be considered more than 40% of the basic per capita payment in deprived areas	Per capita	• Losing 20% of the basic payment if you get 50 negative marks (deduction from salary) • Contract termination if 100 or more warning points are obtained during a contract period (warning points)	Monthly and quarterly
ROSP	Preventive care: 3-5 absolute goals per evaluation • Use of services: absolute purpose	Just encouragement	• "A small percentage of the income of general practitioners" • Prevention (registered patients): 6.86 € • Use of services: 18-20 € depending on the rurality of the area	Fee for service	Preventive care: fixed payment per patient, fixed additional payment for achievement of each criterion • Utilization of services: If the provision of services reaches the required number of patients, the payment is fixed	Yearly
P&FM-P4P	• Increasing the quality of treatment • Increasing patients' satisfaction with doctor's performance • Increase assistant performance score	Both encouragement and punishment	• Up to 30% of the income of the professional component	Fee for service	• The treatment quality score is a score between 0 and 100, which is measured individually at least once every 3 months • The level of patients' satisfaction with the doctor's performance is measured based on the standard patient satisfaction measurement form	It is calculated monthly and quarterly

than 131 measures in about 30 areas. Despite this, there is mixed evidence of teaching to the test in the QOF. One study showed neither deterioration nor improvement in unrewarded conditions [79].

Despite the NHI-P4P design seems to be lacking in several respects, several studies have found positive effects of this program [73, 80]. This may seem surprising, but the shortcomings in the design of mainly to aspects that mitigate undesired behavior, including a relatively narrow definition of performance (concern about teaching to the test), limited provider involvement in design (provider support unlikely) and lack of risk adjustment for outcomes (incentives for selection). Unlike most of the programs reviewed in this study, the P&FM-P4P program focuses on all medical groups that contract with insurance organizations. Like most programs, it also focuses on quality aspects of patient care.

Although the data provide limited information on the use of risk-mitigating measures, the results raise doubts about whether differences in (patient) risk are sufficiently equalized, especially in NHI-P4P, FM PBC, ROSP. In view of the relatively large payments in FM PBC and uncertain financial consequences in ERGOV, concerns about teaching to the test are particularly large in these programs. In ERGOV, rewards for high performers are financed by abuse for low performers. Although this contributes to financial sustainability, it may increase the

incentives for gaming theory [54]. To avoid this, clinics are required to supply data via an online tool that enables checks and auditing.

The negative incentives should be taken seriously given evidence that providers do indeed respond to incentives [81]. Many current P4P-programs have shortcomings with respect to design elements that relate to the prevention of undesired behavior (specifically risk selection and teaching to the test), and there is a great diversity in the use of risk-mitigating measures. This shows that buyers though clearly concerned about them, are unsure about how to effectively prevent undesired effects. Therefore because such effects can potentially undermine the entire program, more insight is needed to prevent them. For example, research should continue to focus on developing sufficient risk adjustment than can be applied transparently in practice and on the drawbacks and merits of potentially viable alternatives or supplements such as exception reporting. Second, if P4P is to help to improving patient outcomes, payment allocation must be based on timely, accurate, and reliable performance data. Many shortcomings in the design of current programs, including small measure sets, low payment frequency, lack of risk adjustment, and limited use of outcomes, can be traced back to a lack of data. Efforts should be focused on developing methods for recording, extracting, and processing patient-level data, and the merits of information technology for these purposes should be explored further. MacDonald and Rowland In their study (2009), found that "requiring data entry into the patient's electronic medical record" was intended to reduce eye contact and increase time spent collecting data [82].

Third, breakthrough improvements require alignment of incentives and coordination across disciplines for all providers in the continuum of care. Current programs focus too much on a type of provider (physician groups in primary care) and specific sector. For example, ROSP and PIN specifically target physicians, and payments are mostly made annually [61]. In FM PBC, even though the bonus amount is given to doctors and nurses, the main focus is on the GP program, also the remuneration is paid directly to the family doctors, and then the family doctor decides to divide the amount among other service providers [58]. Consistent incentives require strategies to facilitate inclusion of small practices (*e.g.,* developing methods for collecting performance data) as well as incorporating incentives that encourage coordination. Forms of prospective payment like bundled case rates and customized IT will prove vital in attaining these goals. If structured around patients rather than providers, prospective payment with performance-based elements can both reward providers for effectively coordinating care and reduce the problem of overuse of low-value services [83].

Fourth, it is very important that programs are evaluated using convincing control groups. Of the identified programs, only five have been evaluated, and often only partially. Therefore, studies should not only assess effectiveness but also include assessments of adverse effects and the impact of specific design elements. This not only provides insight in which areas need modification, but also important lessons about program design.

There are two dominant strategies for calculating indicators and paying bonuses among the studied programs, which are "achieving a certain amount of the measures" and "increasing the measures". In the NHI-P4P, ERGOV, FM PBC, ROSP programs, a minimum standard is considered for each measures, and in the payment formula, the individual's performance score is calculated based on the individual's achievement of the standards. On the other hand, in QOF, AQ, IHA-P4P, PIN, P&FM-P4P programs, if the person's performance is more than the target amount, it will be added to his overall score. Applying both types of incentive payment strategies to providers has the potential to improve the performance of service providers with both below-standard and above-standard performance. A study by Ejkanar et al. showed that rewarding providers based on their achievement of predetermined goals may reward high performers, while the greatest improvement was seen in the lowest performing physicians. This means that physicians who had the poorest performance at the beginning of the program achieved the greatest improvement in performance compared to physicians whose performance was above the target level at the beginning of the program [84]. Some critics have also argued that the use of a certain limit of the index cannot motivate providers whose performance is higher than the target level at the beginning of the program. On the other hand, other providers who previously performed poorly have less motivation to improve their performance. Because according to them, the goals are very difficult to achieve [85].

Conclusions

P4P is now widely being used in many healthcare systems and there are no signs that this will change in the near future. However, current evidence suggests that designing an effective P4P-program is a very complex task. Given the limited knowledge about "what works" in P4P, it may not be too surprising that current program design seems to be lacking in several respects and that buyers struggle with developing effective programs. To get the most out of P4P, well-conducted assessments are critical for generating the information needed for fine-tuning P4P to the specific implementation settings. In particular, empirical research is needed to examine the impact of specific design choices in specific settings, as well as insight in the perverse incentives of P4P and how these can be prevented. In parallel, if P4P is to help improve patient outcomes, efforts should be focused on creating easy-to-use and reliable methods for generating comprehensive patient-level (performance) data.

Ethical Approval

This study was approved by the Ethics Committee of Iran University of Medical Sciences (code: IR.IUMS.REC.1399.407).

Funding/Support

This study was part of a PhD thesis supported by the School of Health Management and Information Sciences, Iran University of Medical Sciences, Tehran, Iran (IUMS/SHMIS-98-3-37-16007).

Conflict of Interests statement

There are no conflicts of interest.

Authors' contributions

Study concept and design: AA and MEE and HP. Analysis of data: HP and MEE. Drafting of the manuscript: MEE and HP. Performed a search of the literature: HP, MEE, MM**.** Critical revision of the manuscript: MB and HAG and JA. Editing: MB, MM. All authors have read and approved the latest version of the paper for publication.

References

[1] Benavent J, Juan C, Clos J, Sequeira E, Gimferrer N, Vilaseca J. Using pay-for-performance to introduce changes in primary healthcare centres in Spain: first year results. Qual Prim Care 2009;17:123-31.

[2] 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;10:1-13. https://doi.org/10.1186/1472-6963-10-247.

[3] Raeisi P, Alikhani M, Mobinizadeh M. Fee for Service (FFS) payment on the basis of performance in Hasheminejad Hospital. Tehran 2010.

[4] Jahani F, Farazi A, Rafiei M, Jadidi R, Anbari Z. Job satisfaction and its related factors among hospital staff in Arak in 2009. J Arak Uni Med Sci 2010;13:32-9.

[5] Gemmell I, Campbell S, Hann M, Sibbald B. Assessing workload in general practice in England before and after the introduction of the pay-for-performance contract. J Adv Nurs 2009;65:509-15. https://doi.org/10.1111/j.1365-2648.2008.04902.x.

[6] Herzer KR, Pronovost PJ. Physician motivation: listening to what pay for performance programs and quality improvement collaboratives are telling us. Jt Comm J Qual Patient Saf 2015;41:522-8. https://doi.org/10.1016/s1553-725041069-4.

[7] Azar FE, Arab M, Badloo M, Rezaei, E. Examining the relationship between Payment based on new manegerial approachfor hospital administration, Expectancy theory and job satisfaction in selected hospital staff of Tehran university of medical sciences in 2013. J Hosp 2015;13:123-32.

[8] Janati A, Valizadeh S, Jafarabadi MA. Development of financial indicators of hospital performance. J Clin Res Govern 2014;3:92-8. https://doi.org/10.13183/JCRG.V3I0.121.

[9] Medina Lara A, Walley T. Health Economics for Developing Countries. A Practical Guide. In: Witter S, Ensor T, Thompson R, Jowett M, eds. London-Oxford: Macmillan Education 2000. https://doi.org/10.1016/S0035-920390355-4.

[10] Roberts MJ, Hsiao W, Berman P, Reich MR. Getting health reform right: a guide to improving performance and equity. New York 2008. https://doi.org/10.1093/acprof:oso/9780195371505.001.0001.

[11] Movahedi M. Pay-per-performance model with an emphasis on socio-cultural characteristics of Iran (Case study: the staff of universities in Tehran city). Tehran: Allameh Tabatabai University 2012.

[12] De Bruin SR, Baan CA, Struijs JN. Pay-for-performance in disease management: a systematic review of the literature. BMC Health Serv Res 2011;11:1-14. https://doi.org/10.1186/1472-6963-11-272.

[13] Aryankhesal A, Behzadifar M, Bakhtiari A, Shahabi S, Azari S, Darvishi Teli B, Rezapour A, Ehsanzadeh SJ, Behzadifar M. Exploring the landscape of health technology assessment in Iran: perspectives from stakeholders on needs, demand and supply. Health Res Policy Syst 2024;22:11. https://doi.org/10.1186/s12961-023-01097-0.

[14] James J. Pay-for-performance. Health Affairs 2012;34:1-6.

[15] Gillam SJ, Siriwardena AN, Steel N. Pay-for-performance in the United Kingdom: impact of the quality and outcomes framework – a systematic review. Ann Fam Med 2012;10:461-8. https://doi.org/10.1370/afm.1377

[16] Piekkola H. Performance-related pay and firm performance in Finland. Int J Manpow 2005;26:619-35. https://doi.org/10.1108/01437720510628103.

[17] Lohmann J, Houlfort N, De Allegri M. Crowding out or no crowding out? A Self-Determination Theory approach to health worker motivation in performance-based financing. Soc Sci Med 2016;169:1-8. https://doi.org/10.1016/j.socscimed.2016.09.006

[18] Chenot R. Pay for performance in dental care: A systematic narrative review of quality P4P models in dental care. ZEFQ 2017;127:42-55. https://doi.org/10.1016/j.zefq.2017.06.001.

[19] Young SG, Gruca TS, Nelson GC. Impact of nonphysician providers on spatial accessibility to primary care in Iowa. Health Serv Res 2020;55:476-85. https://doi.org/10.1111/1475-6773.13280.

[20] Kondo K, Damberg C, Mendelson A, Motu'apuaka M, Freeman M, O'Neil M, Relevo R, Kansagara D. Understanding the Intervention and Implementation Factors Associated with Benefits and Harms of Pay for Performance Programs in Healthcare [Internet]. Washington, DC: Department of Veterans Affairs (US) 2015.

[21] Raiisi P, Alikhani M, Mobinizadeh M. Pay for performance in Shahid hasheminezhad hospital. Health System 2010;2:27-36.

[22] Sura A, Shah NR. Pay-for-performance initiatives: modest benefits for improving healthcare quality. Am Health Drug Benefits 2010;3:135-42.

[23] Doran T, Fullwood C, Gravelle H, Reeves D, Kontopantelis E, Hiroeh U, et al. Pay-for-performance programs in family practices in the United Kingdom. N Engl J Med 2006;355:375-84. https://doi.org/10.1056/NEJMsa055505.

[24] Soucat A, Dale E, Mathauer I, Kutzin J. Pay-for-performance debate: not seeing the forest for the trees. Health Systems & Reform 2017;3:74-9. https://doi.org/10.1080/23288604.2017.1302902.

[25] Kaptanoğlu AY. Performance based supplementary payment systems in Istanbul public hospitals. J High Educ Sci 2013:128-32. https://doi.org/10.5961/jhes.2013.067

[26] Eldridge C, Palmer N. Performance-based payment: some reflections on the discourse, evidence and unanswered questions. Health Policy Plan 2009;24:160-6. https://doi.org/10.1093/heapol/czp002.

[27] Frakt AB, Jha AK. Face the facts: we need to change the way we do pay for performance. Ann Intern Med 2018;168:291-2. https://doi.org/10.7326/M17-3005.

[28] Greengarten M, Hundert M. Individual pay-for-performance in Canadian healthcare organizations. Healthc Pap 2006;6:57-61; discussion 72. https://doi.org/10.12927/hcpap18266.

[29] Alshamsan R, Majeed A, Ashworth M, Car J, Millett C. Impact of pay for performance on inequalities in health care: systematic review. J Health Serv Res Pol 2010;15:178-84. https://doi.org/10.1258/jhsrp.2010.009113.

[30] Sutton M, Nikolova S, Boaden R, Lester H, McDonald R, Roland M. Reduced mortality with hospital pay for performance in England. N Engl J Med 2012;367:1821-8. https://doi.org/10.1056/NEJMsa1114951.

[31] Eichler R, Auxila P, Pollock J. Performance-based payment to improve the impact of health services: evidence from Haiti. World Bank Institute Online Journal 2001;1-11.

[32] Weibel A, Rost K, Osterloh M. Pay for performance in the public sector – Benefits and (hidden) costs. J Public Adm Res Theor 2010;20:387-412.

[33] Kuvaas B, Buch R, Dysvik A, editors. Individual variable pay for performance, incentive effects, and employee motivation. Annual meeting of the Academy of Managemen, Chicago, USA, 2018.

[34] Petersen LA, Woodard LD, Urech T, Daw C, Sookanan S. Does pay-for-performance improve the quality of health care? Ann Intern Med 2006;145:265-72. https://doi.org/10.7326/0003-4819-145-4-200608150-00006.

[35] Lindenauer PK, Remus D, Roman S, Rothberg MB, Benjamin EM, Ma A, et al. Public reporting and pay for performance in hospital quality improvement. N Engl J Med 2007;356:486-96. https://doi.org/10.1056/NEJMsa064964.

[36] Werner RM, Kolstad JT, Stuart EA, Polsky D. The Effect Of Pay-For-Performance In Hospitals: Lessons For Quality Improvement. Health Affairs 2011;30:690-8. https://doi.org/10.1377/hlthaff.2010.1277.

[37] Greenberg JO, Dudley JC, Ferris TG. Engaging specialists in performance-incentive programs. N Engl J Med 2010;362:1558-60. https://doi.org/10.1056/NEJMp1000650.

[38] Martin Roland D. Linking physicians' pay to the quality of care – a major experiment in the United Kingdom. N Engl J Med 2004;351:1448-54. https://doi.org/10.1056/NEJMhpr041294.

[39] West NN. A North West health system approach to advancing quality. Manchester: NHS North West Strategic Health Authority 2008.

[40] Doran T, Kontopantelis E, Reeves D, Sutton M, Ryan AM. Setting performance targets in pay for performance programmes: what can we learn from QOF? BMJ 2014;348:g1595.

[41] Kristensen SR, Meacock R, Turner AJ, Boaden R, McDonald R, Roland M, et al. Long-term effect of hospital pay for performance on mortality in England. N Engl J Med 2014;371:540-8. https://doi.org/10.1056/NEJMoa1400962.

[42] Chang H, ed. Quality-Based Payment – Taiwan's Experience. Annual Research Meeting 2004.

[43] Chang RE, Lin SP, Aron DC. A pay-for-performance program in Taiwan improved care for some diabetes patients, but doctors may have excluded sicker ones. Health Affairs 2012;31:93-102.

PMID: 22232099 https://doi.org/10.1377/hlthaff.2010.0402.

[44] Chen TT, Chung KP, Lin IC, Lai MS. 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 Serv Res 2011;46:47-60. https://doi.org/10.1111/j.1475-6773.2010.01182.x.

[45] Kuo RN, Chung K-P, Lai M-S. Effect of the pay-for-performance program for breast cancer care in Taiwan. JCO Oncol Pract 2011;7:e8s-15. https://doi.org/10.1200/JOP.2011.000314.

[46] Tsai W-C, Kung P-T, Khan M, Campbell C, Yang W-T, Lee T-F, et al. Effects of pay-for-performance system on tuberculosis default cases control and treatment in Taiwan. J Infect 2010;61:235-43. https://doi.org/10.1016/j.jinf.2010.06.016.

[47] Cashin C, Chi Y-L, Smith PC, Borowitz M, Thomson S, eds. Paying for performance in health care: implications for health system performance and accountability. McGraw-Hill Education (UK) 2014.

[48] Robinson JC, Williams T, Yanagihara D. Measurement of and reward for efficiency in California's pay-for-performance program. Health Affairs 2009;28:1438-47. https://doi.org/10.1377/hlthaff.28.5.1438.

[49] Yegian J, Yanagihara D. Value based pay for performance in California. Integr Healthc Assoc 2013;9:1-5.

[50] Association IH. Advancing quality through collaboration: the California pay-for-performance program. Ind Health Law Rev 2006;3(2).

[51] Katz A, Bogdanovic B, Soodeen R. Physician Integrated Network Baseline Evaluation: Linking Electronic Medical Records and Administrative Data. Winnipeg, MB: Manitoba Centre for Health Policy 2010.

[52] Kreindler SA, Metge C, Struthers A, Harlos K, Charette C, Bapuji S, et al. Primary care reform in Manitoba, Canada, 2011-15: balancing accountability and acceptability. Health Policy 2019;123:532-7. https://doi.org/10.1016/j.healthpol.2019.03.014.

[53] Price M, Singer A, Kim J. Adopting electronic medical records: are they just electronic paper records? CFP 2013;59:e322-9.

[54] Gerdes N, Funke UN, Schüwer U, Kunze H, Walle E, Kleinfeld A, et al. Ergebnisorientierte Vergütung der Rehabilitation nach Schlaganfall-Entwicklungsschritte eines Modellprojekts 2001-2008. Die Rehabilitation 2009;48:190-201. https://doi.org/ 10.1055/s-0029-1231061.

[55] Herzogenaurach F. Qualitätsbericht rehabilitation 2009. Hopfen am See: m&i-Klinikgruppe Enzensberg 2010.

[56] Tatar M, Mollahaliloğlu S, Sahin B, Aydin S, Maresso A, Hernández-Quevedo C. Turkey. Health system review. Health Syst Transit 2011;13:1-186, xiii-xiv.

[57] Turkey - Performance based contracting scheme in family medicine: design and achievements. Washington, DC: World Bank 2013. http://documents.worldbank.org/curated/en/880861468338993763.

[58] Menon R, Nguyen SN, Arur A, Yener AL, Postolovska I. Turkey: Family medicine performance based contracting scheme. Paying for Performance in Health Care 2014:189.

[59] Ozcan YA, Khushalani J. Assessing efficiency of public health and medical care provision in OECD countries after a decade of reform. Cent Eur J Oper Res 2017;25. https://doi.org/10.1007/s10100-016-0440-0.

[60] Atramont A, Coffy S, Sebo S, Debeugny G. Pay for performance scheme for general practitioners in France: results in 2018. Eur J Public Health 2019;29:ckz187-47. https://doi.org/10.1093/eurpub/ckz187.147.

[61] Humbert X, Rabiaza A, Bansard M, Soulie B, Raginel T, Le Bas F. Medical training and pay for performance program. Achievement of pay for performance program. Médecine 2018;14:181-6.

[62] Pomey M-P, Denis J-L, Bernier M, Vergnaud S, Préval J, Saint-Lary O. Innovation in physician remuneration in France: what lessons for Canada? Health Reform Observer - Observatoire des Réformes de Santé 2019;7. https://doi.org/10.13162/hroors.v7i2.3578.

[63] Aryankhesal A, Sheldon TA, Mannion R. Role of pay-for-performance in a hospital performance measurement system: a multiple case study in Iran. Health Policy Plan 2013;28:206-14. https://doi.org/10.1093/heapol/czs055.

[64] Jabbari A, Shaarbafchi Zadeh N, Maddahian B. Identifying Executive Challenges of Performance-Based Payment from Medi-

cal and Educational Hospitals Administrators' Perspective and Offering Solutions in Isfahan (2018). Evid Based Health Policy Manag Econ 2019;3:121-30.

[65] Tabrizi JS, Farahbakhsh M, Iezadi S, Ahari AM. Design and implementation of pay-for-quality in primary healthcare: A case study from Iran. Australas Med J 2017;10:449-59. https://doi.org/10.21767/AMJ.2017.2890.

[66] Taheri G, Gorji HA, Behzadifar M, Rezapour A, Martini M, Sirizi MJ. Strategies for Creating Tax-Based Sustainable Financial Resources in Iran's Health System: a Qualitative Study. J Prev Med Hyg 2024;65:E442-8. https://doi.org/10.15167/2421-4248/jpmh2024.65.3.3379.

[67] Constantinou P, Sicsic J, Franc C. Effect of pay-for-performance on cervical cancer screening participation in France. Int J Health Econ Manag 2017;17:181-201. https://doi.org/10.1007/s10754-016-9207-3.

[68] Mehrotra A, Adams JL, Thomas JW, McGlynn EA. Cost profiles: should the focus be on individual physicians or physician groups? Health Affairs 2010;29:1532-8. https://doi.org/10.1377/hlthaff.2009.1091.

[69] Kaplan SH, Griffith JL, Price LL, Pawlson LG, Greenfield S. Improving the reliability of physician performance assessment: identifying the" physician effect" on quality and creating composite measures. Medical care 2009:378-87. https://doi.org/10.1097/MLR.0b013e31818dce07.

[70] Holmboe ES, Weng W, Arnold GK, Kaplan SH, Normand SL, Greenfield S, et al. The comprehensive care project: measuring physician performance in ambulatory practice. Health Serv Res 2010;45:1912-33. https://doi.org/10.1111/j.1475-6773.2010.01160.x.

[71] Gharibi F, Dadgar E. Pay-for-performance challenges in family physician program. Malays Fam Physician 2020;15:19.

[72] Li J, Hurley J, DeCicca P, Buckley G. Physician response to pay-for-performance: evidence from a natural experiment. Health Econ 2014;23:962-78. https://doi.org/10.1002/hec.2971.

[73] Lee T-T, Cheng S-H, Chen C-C, Lai M-S. A pay-for-performance program for diabetes care in Taiwan: a preliminary assessment. Am J Manag Care 2010;16:65-9.

[74] Aghajani MH, Manavi S, Maher A, Rafiei S, Ayoubian A, Shahrami A, et al. Pay for performance in hospital management: A case study. Int J Healthc Manag 2021;14:484-90. https://doi.org/10.1080/20479700.2019.1664029.

[75] Johnson R, Dinakar C. Pediatric pay-for-performance in asthma: who pays? Curr Allergy Asthma Rep 2010;10:405-10. https://doi.org/10.1007/s11882-010-0140-0.

[76] Eijkenaar F. Pay for performance in health care: an international overview of initiatives. Med Care Res Rev 2012;69:251-76. https://doi.org/10.1177/1077558711432891.

[77] Lee PV, Berenson RA, Tooker J. Payment reform – the need to harmonize approaches in Medicare and the private sector. N Engl J Med 2010;362:3-5. https://doi.org/10.1056/NEJMp0910459.

[78] Doran T, Roland M. Lessons from major initiatives to improve primary care in the United Kingdom. Health Affairs 2010;29:1023-9. https://doi.org/10.1377/hlthaff.2010.0069.

[79] Steel N, Maisey S, Clark A, Fleetcroft R, Howe A. Quality of clinical primary care and targeted incentive payments: an observational study. Br J Gen Pract 2007;57:449-54.

[80] Li YH, Tsai WC, Khan M, Yang WT, Lee TF, Wu YC, et al. The effects of pay-for-performance on tuberculosis treatment in Taiwan. Health Policy Plan 2010;25:334-41. https://doi.org/10.1093/heapol/czq006.

[81] Mullen KJ, Frank RG, Rosenthal MB. Can you get what you pay for? Pay-for-performance and the quality of healthcare providers. Rand J Econ 2010;41:64-91. https://doi.org/10.1111/j.1756-2171.2009.00090.x.

[82] 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. https://doi.org/10.1177/1077558705285291.

[83] Rosenthal MB, Frank RG. What is the empirical basis for paying for quality in health care? Med Care Res Rev 2006;63:135-57. https://doi.org/10.1177/1077558705285291.

[84] Eijkenaar F. Key issues in the design of pay for performance programs. Eur J Health Econ 2013;14:117-31. https://doi.org/10.1007/s10198-011-0347-6.

[85] Chen JY, Kang N, Juarez DT, Hodges KA, Chung RS. Impact of a pay-for-performance program on low performing physicians. J Healthc Qual 2010;32:13-22. https://doi.org/10.1111/j.1945-1474.2009.00059.x.

Received on January 16, 2024. Accepted on December 17, 2024.

Correspondence: Aidin Aryankhesal, Health Management and Economics Research Center, Health Management Research Institute, Iran University of Medical Sciences, Tehran, Iran. E-mail: aryankhesal.a@iums.ac.ir

How to cite this article: Eghbali ME,Pourasghari H, Gorji Abolghasem H, Martini M, Arabloo J, Behzadifar M, Aryankhesal A. Performance-based payment systems for general practitioners and specialists in selected countries: a comparative study J Prev Med Hyg 2025;66:E114-E125. https://doi.org/10.15167/2421-4248/jpmh2025.66.1.3511