Journal of Medicine and Public Health

ISSN: 2766-8355

Review Article

The Impact of Pay-for-Performance Schemes on the Quality of Healthcare in Global Context - Scoping Review

Ahsana Nazish[1], Hemali Jayasekera[2], Omair Ayaz[1], Sana Hyat[1], Wafa Aftab[1] and Paramjit Gill[2]

[1]Department of Community Health Sciences, Aga Khan University, Pakistan

[2]University of Warwick, UK

Abstract

Background: Pay-for-Performance (P4P) schemes provide financial incentives to providers based on goal-directed achievements and have been widely implemented across various settings. The impact of this model on the quality of health care has been uncertain. The purpose of this review is to gather knowledge and experience from existing P4P programs and examine their effect on the quality of healthcare.

Methods: We conducted a scoping review, using data sources; PubMed, CINAHL Plus, Cochrane and ProQuest for a comprehensive literature search from January, 1980 to August, 2021. The extracted data include publication date, setting, design features of pay for performance schemes and their outcomes. A reporting framework and typology were used to categorize P4P design.

Results: Our search yielded 3625 articles of which 94 studies were included. Overall, 75% of the studies were from high and upper-middle-income countries. We found substantial heterogeneity in the design of P4P schemes. The impact of P4P schemes on health care quality showed mixed results that ranged from positive to negative or no effect at all.

Conclusion: The variation in impact can be explained due to the heterogeneity in P4P design and setting where it was executed. Interventions found in one setting may not improve outcomes when implemented in another setting. Therefore, healthcare system readiness should be assessed before the implementation of complex interventions like P4P.

Keywords: Pay for performance; Provider incentives; Impact; Healthcare

Introduction

Pay-for-Performance (P4P), also called Performance-Based Financing (PBF) or Results-Based Financing (RBF) has become a popular method to foster improvement in the quality of health care [1]. P4P programs utilize financial incentives to galvanize improvement in quality, efficiency, outcomes and value of healthcare. Performance metrics are used to monitor quality and outcome [1-3]. This motivates providers to shift focus away from utilization toward patient care [3]. P4P can be implemented at many tiers of healthcare and extensively varies in design. Absolute performance at the provider level may be used by P4P schemes to apply reward-based payment models. Alternatively, P4P programs may consider the relative performance of practitioners to adjust physician income [4]. P4P can also pave the way for healthcare reform such as the adaptation of electronic records, introduction of novel data verification systems and financial decentralization [4,5].

P4P has been well documented in High-Income Countries (HICs) [6-12]. More than half of private Health Maintenance Organizations (HMO) in the United States (US) has adopted the finance-based incentive model [7]. The Centers for Medicare and Medicaid Services (CMS) have also incorporated P4P into Medicare as solicited by the US Congress [7]. In the United Kingdom (UK), the Quality and Outcomes Framework (QOF) for family practitioners was implemented in 2004 [8]. QOF is designed to remunerate general practices for providing quality care to the patients and help fund work to further improve the quality of healthcare delivered [9] The Australian Government introduced the Practice Incentives Program (PIP) in 2001 to help curb preventable complications in diabetic and asthmatic patients [10]. Furthermore, P4P has also been heavily investigated in terms of efficacy for implementation in Canada [12]. These programs encompass sizable portions of the population in each setting and have resulted in mixed outcomes [7-12].

Data on P4P programs implemented in Low-and-Middle-Income Countries (LMICs) has shown varied success [13-18]. Rwanda was the first of many LMICs to implement these schemes, where P4P showed improvement in the use and quality of Maternal and Child Health (MCH) [19]. A study in China showed that P4P improved antibiotic prescription practices and reduced spending in primary care [17]. However, not all outcomes of P4P have proven to be beneficial [20]. A cluster-randomized trial in Afghanistan showed a minimal effect in improving MCH [21]. P4P schemes have either shown no effect or even a decrease in quality of care in other settings [22,23]. This could be sourced to the differences in design features of these models. Although there is an abundance of literature evaluating P4P outcomes, there is a scarcity of studies that investigate the design of these schemes [24].

Variation in the success of P4P could be explained by differences in design features of P4P research studies and programs. Poor design

Citation: Nazish A, Jayasekera H, Ayaz O, Hyat S, Aftab W, Gill P. The Impact of Pay-for-Performance Schemes on the Quality of Healthcare in Global Context - Scoping Review. J Med Public Health. 2022;3(6):1050.

Publisher Name: Medtext Publications LLC

Manuscript compiled: Nov 23th, 2022

***Corresponding author:** Ahsana Nazish, Department of Community Health Sciences, Aga Khan University, Karachi, Pakistan, E-mail: ahsanazish80@gmail.com

is partly due to the lack of a theoretical/conceptual framework as to why the particular design used is likely to influence behavior. A recent systematic review by Kovacs et al. [5] introduced a typology with 5 parameters to classify and describe the design of P4P schemes in LMICs [25]. In this study, we have adopted a theoretically informed reporting framework and typology for categorizing and analyzing designs of healthcare pay-for-performance schemes introduced by Ogundeji et al. [24].

The aim of this review is to examine empirical literature available globally about the "high-level design features" of pay-for-performance schemes and opportunities for improving reporting of design features. The objective of this review is:

- To categorize the key design features of P4P schemes in the healthcare setting by using a theoretically informed reporting framework.
- Examine the outcomes reported in pay-for-performance schemes on quality of care.

Materials and Methods

We used the Arksey and O'Malley methodological framework for scoping review and adhere to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-SR) guidelines [26]. To evaluate and compare the effect of design on outcomes of P4P schemes developed across the world, we used a theoretically informed reporting HISReF framework [24].

Literature search

A search strategy was developed in consultation with an experienced librarian. To capture a maximum number of studies in our review, we included research published and unpublished over the last 40 years (i.e., from January 1980 to August 2021). The databases for literature search include PubMed, CINAHL Plus, Cochrane and ProQuest. The titles and abstracts of relevant articles, and the index terms used to describe the articles were extracted. A second search was then performed using the terms identified from the initial review. The search terms used were: (Pay-for-performance OR pay for performance OR P4P) and (incentive OR bonus OR reward OR compensation OR performance-based financing) AND (programs OR programme OR health care providers OR doctor OR nurse OR community health worker OR physician OR surgeon) and (quality of health care OR quality of care OR communities OR hospitals OR primary healthcare OR secondary healthcare OR tertiary healthcare OR hospital OR clinic) References from retrieved articles were then screened for additional research for the final stage of the process (Annex 1).

Study selection

A two-stage screening process that comprises of title and abstract screening, followed by full review of selected articles to determine eligibility for inclusion. Every article was reviewed independently by two reviewers (AN and HJ). The review included full spectrum of study designs i.e., quantitative, qualitative or mixed methods, which used primary data analysis and reported on evaluation of P4P programmes. We agreed to exclude commentaries, book reviews, conference abstracts and editorial articles. Articles that did not have details on the design features (incentive recipients, type of incentives, size of incentive and perceived risk) of the P4P model were also excluded.

Data extraction and collation

The lists of all citations retrieved were managed through EndNote X9. All unnecessary and duplicate records were eliminated. Two independent reviewers screened the title and abstract of each article to determine its relevance. The full text of selected citations was reviewed in detail against the inclusion criteria by the same independent reviewers. Reasons for exclusion were also documented. A standardized charting form was developed to categorize data. It included information on publication, authors and study design features (Annex 2). Components of design such as incentive recipients, types of incentives, size of incentives, perceived risk, and outcomes were captured (Figure 1).

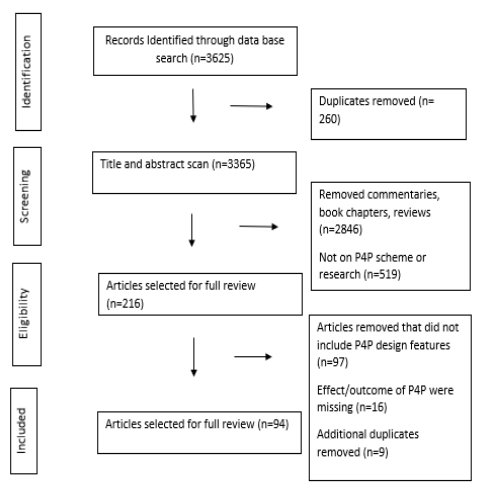

Figure 1: Study search and selection.

Results

Using the search strategy described previously, a total of 3625 publications were identified from selected databases. 260 duplicates were removed. Ninety four studies were eventually filtered in for further review at the end of the screening process, as mentioned in the flowchart (Figure 1). The characteristics of included studies for review are listed in Table 1. Thirty three included studies were conducted in North America (31 in the US and 2 in Canada), 24 in Asia, 20 in Africa, 4 in Australia and 7 in Europe (3 in France, 2 in Germany, 1 in Spain and 1 in Portugal). In terms of study design, 85 studies were identified as quantitative. Twenty eight of those were interventional, randomized and quasi-experimental. Eighteen studies used an observational cohort design. Twelve were cross-sectional studies and 11 were longitudinal studies. Seven used comparison groups and 8 were descriptive. Three studies were qualitative and 9 were mixed (qualitative and quantitative). Majority of the Studies are mostly from HICs (n=64), followed by LMICs (n=24) and the rest were from upper-middle-income countries (UMICs, n=8) (Table 1).

Table 2 provides information on the design features of the included studies.

1. Incentive recipients (individuals or group): Pay for performance schemes involve payment of financial incentives to individual health care providers or groups such as institutions and clinical teams. Table 2 presents, 72 included

Table 1: Characteristics of included studies for the P4P scoping review.

Included Studies	N
Total	94
Methodology	
Quantitative	85
Qualitative	3
Mixed	9
Country	
United States	31
United Kingdom	6
Europe	9
Africa	20
Canada	2
Australia	4
Brazil	1
China	3
Taiwan	16
Afghanistan	2
Philippines	1
Korea	1

Table 2: Design features of included studies.

Design Features	N
Total	94
Who is Incentivised	
Individuals	23
Groups	69
Both	2
Type of Incentive	
Bonus	83
Penalty	1
Both	7
Size of Incentive	
Small	8
Large	55
Not mentioned	31
Risk of Not Earning Incentive	
Low	44
High	12
Not Mentioned	38

studies incentivized groups, 23 incentivized individuals and 2 incentivized both.

2. Form of incentive (bonus or fine): The 2 forms of financial incentives used in P4P schemes are fines and bonuses. Table 2 shows, 85 included studies used bonuses, 8 studies used both bonuses and fines, 1 study used fines and 3 studies did not mention the type of incentive used.

3. Size of incentive (small or large): The amount of money rewarded/penalized in relation to the provider's salary, budget of health institution or payment for a health service. We classified this into 2 categories: small (incentive < 5%) and large (incentive ≥ 5%). 55 included studies had a large incentive, 8 had a small incentive and 31 did not report the size of the incentive.

4. Perceived risk of not earning the incentive (high or low): The fourth domain proposed by Ogundeji et al. [24] is based on risk aversion theory. It explains the behavior of individuals when exposed to uncertainty or risk. Individuals are less likely to change their behavior when the risk of not getting an expected reward is high. Time till receiving payment after achieving targets (time lag), domain of performance measure and performance measure (payment scale) were used to determine the risk. Forty four included studies were low risk, 12 were high risk and 38 did not have enough information for classification (Table 2).

Reported outcomes of P4P schemes

The effects of P4P interventions in the included studies ranged from positive (70%) to negative or non-effectual (30%). Negative outcomes indicate no improvement and not decline in quality of care. Performance of P4P programs were measured using service utility, quality of care and patient outcomes. LMICs commonly use P4P schemes to achieve unmet Millennium Development Goals 4 and 5 for Maternal and Child Health (MCH) [25]. The effects of P4P on MCH services were greater in some settings where there was an increase in institutional deliveries and improvement in quality of maternal and neonatal care [26]. In Tanzania, P4P helped to improve the utilization of nearest health centers for pregnant women [27]. Studies conducted in Burundi and Rwanda showed improvement in the use and quality of MCH services [28,29]. Moreover, patient satisfaction for MCH services particularly improved by incentivizing individual providers [30]. P4P schemes are increasingly being implemented in LMICs for quality improvement in care associated for Non-Communicable Diseases (NCDs) as well. P4P programs targeting chronic illnesses like Cardiovascular Disease (CVD), Diabetes Mellitus (DM) and Hypertension (HTN) have resulted in better compliance and improvement of indicators. Ten studies targeting diabetic patients showed a significant reduction in mortality, increase in patient awareness and decrease in hospitalization for complications [31-40].

A study conducted in Sweden revealed that P4P increased compliance to drug treatment guidelines for HTN; however, the effect was strong for private providers [41]. A reduction in default rate for patients was reported by incentivizing field case managers [42]. Payment enhancement programs have shown remarkable results in decreasing the long-term default rate of patients with chronic illnesses and motivating health workers for better performance [43]. There was a reduction in the cost of care while maintaining its quality when physicians were incentivized individually [44]. P4P schemes implemented in Brazil decreased socioeconomic inequities and improved performance of primary care providers over the course of program. This was potentially due to design features that adjusted for financial payments and socioeconomic inequalities [45].

However, in a few studies, improvement was observed in some areas while there was no effect on other indicators. For example, Rudasingwa et al. [46] reported improvement in performance quality scores of care management and clinical health services (outpatient, prenatal and maternal), but noticed no effect on support services (lab work and material management). P4P supply-side incentive programs showed that improvement in quality of treatment received by underserved children was conditional on care-seeking behavior of parents, but incentivization did not impact the propensity to seek care. These findings provide further evidence that PBF incentivizes the critical role that staff play in assuring quality of services but does little to influence patient demand for these services [47].

There were also studies that show that P4P did not prove to be beneficial in terms of providing desired results. A study conducted in 5 African countries on neonatal health showed that no impact of P4P on early neonatal health outcomes and had little to no effect on utilization and quality of neonatal care [48]. Similarly, another study conducted in Burkina Faso revealed that absence of significant effects by P4P schemes on effective coverage of curative child health services

in low-income settings [49]. There is also concern that long-term impact of P4P initiatives is not sustainable. The initial improvement in quality becomes complacent over time. This was evident in one study conducted in the UK that reported lack of sustainability in improvement of patient outcomes [50]. A program for cataract surgery patients in a high-income setting failed to show positive long-term effects from a P4P program implemented to improve quality of care, and the authors suggested that the program should be designed carefully in order to maximize possible effects [51]. In another setting where P4P was conducted to reduce expenditure on drugs, it consequently resulted in an increase of expenditures on non-drug related services. This was attributed to failure is design of the P4P scheme. Evidence from a study in Rwanda showed that P4P was accompanied by considerable secondary effects. Management of rewards between institutional, departmental and individuals should be regulated. Otherwise, individuals might find themselves under institutional pressure to report the expected results; whether they are achieved or not [52].

Discussion

This review examined the design of P4P schemes in healthcare across the globe and identified 94 P4P schemes that were implemented in 23 countries. We found substantial heterogeneity in the design of P4P schemes which can be explained by an array of choice that goes into program design. Results strongly suggest a variety in terms of who receives incentives, size and type of incentives, and risk associated with them. This supports the idea that there is no standard P4P design that can be applied in a health care setting. Group incentives are preferred over individual incentives as performance issues are improved most efficiently through group effort rather than individual effort [53-57]. However, it is important to assess how incentives are distributed between group members [58]. Bonuses were found to be the most common form of incentives used in P4P programs, although, some authors believe that fines could be more sustainable in terms of cost [56]. Provider performance could also improve if they think that they might earn less. However, research shows that programs incorporating losses are perceived as unfair and can result in negative responses from the incentivized. In practice, the use of P4P programs enabling fines has been on the decline [59].

Information on important design features like size of incentives (small or large), perceived risk of not earning incentives (payment timeline, relative or absolute payment, and clinicians' control over receiving payment) is unreported in most studies. To begin to understand what schemes work in which settings, it is important to have information on the design of schemes available. Although researchers have proposed typology to analyze P4P design, relevant information on key design elements remains unclear. We found that the HISReF reporting framework useful in assessing P4P schemes, yet information to use the tool appropriately and completely was absent in studies. Given the multiplicity in P4P design, we suggest that authors should fully describe design elements in future studies. Information on design will be helpful for policymakers who are desire to use incentives to improve provider performance.

We also found that the impact of P4P schemes shows mixed results. These range from positive to negative, or sometimes have no effect. For example, in a study by Rudasingwa et al. [60] that was launched with the assistance of the Ministry of Health in Burundi, it was concluded that the PBF scheme shows improvement of the performance quality scores of healthcare management and services. Outpatient care, prenatal care, and maternal care all exhibited a meaningful trend. However, the same financial incentives had no effect on clinical support services such as laboratory facilities, pharmaceutical management and healthcare materials management. During the same year, Bonfrer et al. [19] studied and assessed the same program in Burundi and concluded that, "the introduction of performance-based financing in Burundi improved the use of maternal care services and the quality of health care services during the period 2006-10." More recently, A. Gage and S. Bauhoff investigated P4P programs in 5 African countries (including Burundi) and stated the schemes had no impact on health outcomes and had variable effects on utilization of healthcare services. Furthermore, effects of P4P can vary widely due to several contextual factors such as the magnitude of incentives, length of intervention, how goals are measured, etc. Information regarding indicators/metrics (structure, process, outcome) used to stipulate conditions for rewards was also missing in some studies. Only a few studies declared the process involves indicator development. This information is crucial in understanding the impact of programs. There is a lack of evidence on unintentional consequences of P4P intervention as well. Despite this missing information, new programs continue to be brought forward for development; further risking financial resources.

Strengths and limitations of the scoping review

We were able to examine the most recent literature yet did not examine the methodological quality of individual studies. Since this was a scoping review, we were aiming to summarize the entire volume of available evidence within a restricted time framework.

Since 67% (n=65) of the studies were conducted in HICs, the findings were not generalized due to variations in the service provision, size of the incentives, the incentivized party and the duration of the program amongst other factors. The reviewers had to discard most of the studies conducted in LMICs due to under-reporting or lack of data. There was considerable variation in the way which P4P schemes were designed, especially in terms of payments.

Conclusion

We found heterogeneity in the design of P4P programs due to different typologies used by researchers. Authors found the HISReF reporting framework as a useful tool to describe and categorize P4P schemes. However, the information to appropriately use this tool was absent in many studies. This variation in the impact of P4P schemes can be explained due to variability in the design and setting where it was executed. Interventions found in one setting may not improve outcomes when implemented in another setting. This is particularly true if healthcare system elements required to support implementation are not considered. Authors recommend that careful healthcare system readiness be assessed before implementing complex interventions. Therefore, integrating the health system readiness assessment as a precursor to the implementation phase of pilot P4P programs can anticipate shortcomings and help adopt new measures that will enhance effectiveness and uptake.

Ethics and Dissemination

This proposal reports a rigorous, comprehensive and transparent methodology. The review will contribute to the advancement of research on this subject and comment on the maturity of the body of literature by identifying gaps in knowledge. As the scoping review methodology consists of collecting data and reviewing publicly available material, this study does not require ethics approval.

References

1. Kondo KK, Damberg CL, Mendelson A, Motuapuaka M, Freeman M, O'Neil M, et al. Implementation Processes and Pay for Performance in Healthcare: A Systematic Review. J Gen Intern Med. 2016;31 Suppl 1(Suppl 1):61-9.
2. Institute of Medicine (US) Committee on Quality of Health Care in America. Crossing the quality chasm: A new health system for the 21st Century. Washington (DC): National Academies Press (US); 2001.
3. Kahn JM, Scales DC, Au DH, Carson SS, Curtis JR, Dudley AR, et al. An official American Thoracic Society policy statement: pay-for-performance in pulmonary, critical care, and sleep medicine Am J Respir Crit Care Med. 2010;182(7):752-61.
4. Eichler R, Levine R, and the Performance-Based Incentives Working Group. Performance incentives for global health: potential and pitfalls. June 15, 2009.
5. Singh NS, Kovacs RJ, Cassidy R, Kristensen SR, Borghi J, Brown GW. A realist review to assess for whom, under what conditions and how pay for performance programmes work in low- and middle-income countries. Soc Sci Med. 2021;270:113624.
6. Vellez M. How incentive payments support Universal Health Coverage, in theory and in practice [Internet]. World Bank Group; 2019.
7. Epstein AM. Pay for performance at the tipping point. N Engl J Med. 2007;356(5):515-7.
8. Campbell S, Reeves D, Kontopantelis E, Middleton E, Sibbald B, Roland M. Quality of primary care in England with the introduction of pay for performance. N Engl J Med. 2007;357:181-90.
9. Scott IA. Pay for performance in health care: strategic issues for Australian experiments. Med J Aust. 2007;187(1):31-5.
10. Greene J. An examination of pay-for-performance in general practice in Australia. Health Serv Res. 2013;48(4):1415-32.
11. Millett C, Gray J, Saxena S, Netuveli G, Majeed A. Impact of a pay-for- performance incentive on support for smoking cessation and on smoking prevalence among people with diabetes. CMAJ. 2007;176(12):1705-10.
12. Kyeremanteng K, Robidoux R, D'Egidio G, Fernando SM, Neilipovitz D. An Analysis of Pay-for-Performance Schemes and Their Potential Impacts on Health Systems and Outcomes for Patients. Crit Care Res Pract. 2019;2019:8943972.
13. Montagu D, Yamey G. Pay-for-performance and the millennium development goals. Lancet. 2011;377(9775):1383-5.
14. Miller G, Babiarz KS. Pay-for-performance incentives in low- and middle- income country health programs. NBER. 2013.
15. Das A, Gopalan SS, Chandramohan D. Effect of pay for performance to improve quality of maternal and child care in low- and middle-income countries: a systematic review. BMC Public Health. 2016;16(1):321.
16. Witter S, Fretheim A, Kessy FL, Lindahl AK. Paying for performance to improve the delivery of health interventions in low- and middle-income countries. Cochrane Database Syst Rev. 2012;2:Cd007899.
17. Kok MC, Dieleman M, Taegtmeyer M, Broerse J, Kane SS, Ormel H, et al. Which intervention design factors influence performance of community health workers in low- and middle-income countries? A systematic review. Health Policy Plan. 2015;30(9):1207-27.
18. Eijkenaar F, Emmert M, Scheppach M, Schöffski O. Effects of pay for performance in health care: a systematic review of systematic reviews. Health Policy. 2013;110(2-3):115-30.
19. Bonfrer I, Soeters R, Van de Poel E, Basenya O, Longin G, van de Looij F, et al. Introduction of performance-based financing in burundi was associated with improvements in care and quality. Health Aff (Millwood). 2014;33(12):2179-87.
20. Snyder L, Neubauer RL. Pay-for-performance principles that promote patient-centered care: an ethics manifesto. Ann Intern Med. 2007;147(11):792-4.
21. Engineer CY, Dale E, Anubhav A, Arunika A, Alonge O, Edward A, et al. Effectiveness of a pay-for-performance intervention to improve maternal and child health services in Afghanistan: a cluster- randomized trial. Int J Epidemiol. 2016;45(2):451-9.
22. Powell-Jackson T, Yip WC-M, Han W. Realigning demand and supply side incentives to improve primary health care seeking in rural China. Health Econ. 2015;24(6):755-72.
23. Huillery E, Seban J. Financial Incentives, Efforts, and Performances in the Health Sector: Experimental Evidence from the Democratic Republic of Congo. Econ Develop Cult Change. 2019:69(3):703235.
24. Ogundeji YK, Sheldon TA, Maynard A. A reporting framework for describing and a typology for categorizing and analyzing the designs of health care pay for performance schemes. BMC Health Serv Res. 2018;18(1):686.
25. Basinga P, Gertler PJ, Binagwaho A, Soucat AL, Sturdy J, Vermeersch CM. Effect on maternal and child health services in Rwanda of payment to primary health-care providers for performance: an impact evaluation. Lancet. 2011;377(9775):1421-8.
26. Rudasingwa M, Soeters R, Bossuyt M. The effect of performance-based financial incentives on improving health care provision in Burundi: a controlled cohort study. Glob J Health Sci. 2015;7(3):15-29.
27. Bezu S, Binyaruka P, Somville V, Mæstad O. Pay-for-performance reduces bypassing of health facilities: Evidence from Tanzania. Soc Sci Med. 2020;268:113551.
28. Bonfrer I, Soeters R, Van de Poel E, Basenya O, Longin G, van de Looij F, et al. Introduction of performance-based financing in Burundi was associated with improvements in care and quality. Health Aff (Millwood). 2014;33(12):2179-87.
29. Bonfrer I, Van de Poel E, Van Doorslaer E. The effects of performance incentives on the utilization and quality of maternal and child care in Burundi. Soc Sci Med. 2014;123:96-104.
30. Tawfiq E, Desai J, Hyslop D. Effects of results-based financing of maternal and child health services on patient satisfaction in Afghanistan. J Health Serv Res Policy. 2019;24(1):4-10.
31. Chen JY, Tian H, Juarez DT, Hodges KA, Brand J, Chung R, et al. The effect of a PPO pay-for-performance program on patients with diabetes. Am J Manag Care. 2010;16(1):e11-9.
32. Gage A, BauhoS. The effects of performance-based financing on neonatal health outcomes in Burundi, Lesotho, Senegal, Zambia and Zimbabwe. Health Policy Plan. 2021;36(3):332-40.
33. Hsieh HM, Lin TH, Lee IC, Huang CJ, Shin SJ, Chiu HC. The association between participation in a pay-for-performance program and macrovascular complications in patients with type 2 diabetes in Taiwan: a nationwide population-based cohort study. Prev Med. 2016;85:53-9.
34. Hsieh HM, Shin SJ, Tsai SL, Chiu HC. Effectiveness of pay-for-performance incentive designs on diabetes care. Med Care. 2016;54(12):1063-9.
35. Kung FP, Tsai CF, Lu CL, Huang LC, Lu CH. Diabetes pay-for-performance program can reduce all-cause mortality in patients with newly diagnosed type 2 diabetes mellitus. Medicine (Baltimore). 2020;99(7):e19139.
36. Lin TY, Chen CY, Huang YT, Ting MK, Huang JC, Hsu KH. The effectiveness of a pay for performance program on diabetes care in Taiwan: a nationwide population-based longitudinal study. Health Policy. 2016;120(11):1313-21.
37. Sheen YJ, Kung PT, Kuo WY, Chiu LT, Tsai WC. Impact of the pay-for-performance program on lower extremity amputations in patients with diabetes in Taiwan. Medicine (Baltimore). 2018;97(41):e12759.
38. Chou CW, Kung PT, Chou WY, Tsai WC. Pay-for-performance programmes reduce stroke risks in patients with type 2 diabetes: a national cohort study. BMJ Open. 2019;9(10):e026626.
39. 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(4):334-41.
40. Hsieh HM, Chiu HC, Lin YT, Shin SJ. A diabetes pay-for-performance program and the competing causes of death among cancer survivors with type 2 diabetes in Taiwan. Int J Qual Health Care. 2017;29(4):512-20.
41. Ellegård LM. Effects of pay-for-performance on prescription of hypertension drugs among public and private primary care providers in Sweden. Int J Health Econ

Manag. 2020;20(3):215-28.

42. Tsai WC, Kung PT, Khan M, Campbell C, Yang WT, Lee TF, et al. Effects of pay-for-performance system on tuberculosis default cases control and treatment in Taiwan. J Infect. 2010;61(3):235-43.

43. Alonge O, Lin S, Igusa T, Peters D. Improving health systems performance in low- and middle-income countries: a system dynamics model of the pay-for-performance initiative in Afghanistan. Health Policy Plan. 2017;32(10):1417-26.

44. Leitman IM, Levin R, Lipp MJ, Sivaprasad L, Karalakulasingam CJ, Bernard DS, et al. Quality and financial outcomes from gainsharing for inpatient admissions: A three-year experience. J Hosp Med. 2010;5(9):501-7.

45. Roberts ET, Zaslavsky AM, Barnett ML, Landon BE, Ding L, McWilliams JM. Assessment of the effect of adjustment for patient characteristics on hospital readmission rates: implications for pay for performance. JAMA Inter Med. 2018;178(11):1498-507.

46. Rudasingwa M, Soeters R, Basenya O. The effect of performance-based financing on maternal healthcare use in Burundi: a two-wave pooled cross-sectional analysis. Glob Health Action. 2017;10(1):1327241.

47. Skiles MP, Curtis SL, Basinga P, Angeles G, Thirumurthy H. The effect of performance-based financing on illness, care-seeking and treatment among children: an impact evaluation in Rwanda. BMC Health Serv Res. 2015;15:375.

48. Gage A, BauhoS. The effects of performance-based financing on neonatal health outcomes in Burundi, Lesotho, Senegal, Zambia and Zimbabwe. Health Policy Plan. 2021;36(3):332-40.

49. Koulidiati JL, De Allegri M, Lohmann J, Hillebrecht M, Kiendrebeogo JA, Hamadou S, et al. Impact of Performance-Based Financing on effective coverage for curative child health services in Burkina Faso: Evidence from a quasi-experimental design. Trop Med Int Health. 2021;26(8):1002-13.

50. 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 Eng J Med. 2014;371(6):540-8.

51. Herbst T, Foerster J, Emmert M. The impact of pay-for-performance on the quality of care in ophthalmology: empirical evidence from Germany. Health Policy. 2018;122(6):667-73.

52. Kalk A, Paul FA, Grabosch E. 'Paying for performance' in Rwanda: does it pay off? Trop Med Int Health. 2010;15(2):182-90.

53. Bertone MP, Lagarde M, Witter S. Performance-based financing in the context of the complex remuneration of health workers: findings from a mixed-method study in rural Sierra Leone. BMC Health Serv Res. 2016;16:286.

54. Chien AT, Eastman D, Li Z, Rosenthal MB. Impact of a pay for performance program to improve diabetes care in the safety net. Prev Med. 2012;55 Suppl:S80-5.

55. Falisse JB, Ndayishimiye J, Kamenyero V, Bossuyt M. Performance-based financing in the context of selective free health-care: an evaluation of its effects on the use of primary health-care services in Burundi using routine data. Health Policy Plan. 2015;30(10):1251-60.

56. Pope GC. Overview of pay for performance models and issues. In: Cromwell J, Trisolini MG, Pope GC, Mitchell JB, Greenwald LM, editors. Pay for performance in health care: methods and approaches. North Carolina: Research Triangle Press; 2011. p. 33-60.

57. 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(1 Pt 1):47-60.

58. An LC, Bluhm JH, Foldes SS, Alesci NL, Klatt CM, Nersesian WS, et al. A randomized trial of a pay-for-performance program targeting clinician referral to a state tobacco quitline. Arch Inter Med. 2008;168(18):1993-9.

59. Grossbart SR. What's the return? Assessing the effect of "pay-for-performance" initiatives on the quality of care delivery. Med Care Res Rev. 2006;63(1 suppl):29S-48S.

60. Rudasingwa M, Soeters R, Basenya O. The effect of performance-based financing on maternal healthcare use in Burundi: a two-wave pooled cross-sectional analysis. Glob Health Action. 2017;10(1):1327241.

Annex 1: Search strategy:

- PubMed search strategy as on October 11, 2021:

  (Pay-for-performance OR pay for performance OR P4P) AND (incentive OR bonus OR reward OR compensation OR performance-based financing) AND (programs OR programme OR health care providers OR doctor OR nurse OR community health worker OR physician OR surgeon) AND (quality of health care OR quality of care OR communities OR hospitals OR primary healthcare OR secondary healthcare OR tertiary healthcare OR hospital OR clinic)

  Period: 01-01-1980 to Oct 11, 2021

  Filters applied: Human, English

  Results: 2788

- EBSCO CINAHL Plus search strategy:

  (Pay-for-performance OR pay for performance OR P4P) AND (incentive OR bonus OR reward OR compensation OR performance-based financing) AND (programs OR programme OR health care providers OR doctor OR nurse OR community health worker OR physician OR surgeon) AND (quality of health care OR quality of care OR communities OR hospitals OR primary healthcare OR secondary healthcare OR tertiary healthcare OR hospital OR clinic)

  Expanders - Apply equivalent subjects

  Search modes - Boolean/Phrase

  Results: 833

- Cochrane data base search strategy:

  (Pay-for-performance OR pay for performance OR P4P) AND (incentive OR bonus OR reward OR compensation OR performance-based financing) AND (programs OR programme OR health care providers OR doctor OR nurse OR community health worker OR physician OR surgeon) AND (quality of health care OR quality of care OR communities OR hospitals OR primary healthcare OR secondary healthcare OR tertiary healthcare OR hospital OR clinic) in Title Abstract Keyword

  Period: 01-01-1980 to Oct 11, 2021

  Results: 27

- ProQuest Dissertations & Theses Database:

  (Pay-for-performance OR pay for performance OR incentive OR bonus OR reward OR P4P OR compensation OR performance based financing) AND (programs OR programme OR health care providers OR doctor OR nurse OR nurse OR community health worker OR physician OR surgeon) AND (quality of health care OR quality of care OR communities OR hospitals OR primary healthcare OR secondary healthcare OR tertiary healthcare OR hospital OR clinic)

  Period: 01-01-1980 to Oct 11, 2021

  Results: 57

Annex 2:

1. Author
2. Study title
3. Journal
4. Study design
5. P4P design feathers

   Item 1- Who received the incentive? Did Individuals or Groups receive the incentive?

   - Individuals: If the incentives are paid directly to individual health workers/clinicians/doctors only
     • If individual health worker/clinician/doctor's income is supplemented as a result of the incentive (e.g., reflected in the rise of personal income) only
   - Groups: If the incentive is paid to a group or an organization in which individual clinicians may or may not benefit from the incentive directly Groups include any of the following
     • Hospital
     • Clinical team
     • General physician (GP) practice
     • NGO
     • Levels of government
     • Faith based organizations
6. Item 2- Type of incentive was the incentive in the form of Fines or Bonuses?
   - Bonus: If incentive is in the form of increase in payments, bonus, gifts etc. with NO fines levied
   - Fines: If the incentive is negative in the form of reduction in expected payments, penalty, punishment etc. In some cases, bonuses may or may not be paid.
7. Item 3- Size of the incentive Was the size of the incentive small or large?
   - Small: If the incentive in the P4P program is smaller than 5% of any one of the following:
     • Salary of individual clinician/health worker/doctor
     • Anticipated payments (to the health facility/hospital/clinical team) such as budgets (total budget or budget for the particular intervention in question), fee for service (FFS) and capitation
   - Large: If the incentive in the P4P program is 5% and above of any one of the following:
     • Salary of individual clinician/health worker/doctor
     • Anticipated payments (to the health facility/hospital/clinical team) such as budgets (total budget or budget for the particular intervention in question), fee for service (FFS) and capitation
8. Item 4- Perceived Risk of not earning the incentive: High risk or low risk? (based on: Timing of payment after achieving targets (time lag), Domain of performance measure, and Performance measure (payment scale)
   - Criteria for judging risk High:

     If the P4P program has 2 or more of the following features
     • If incentive payment (or penalty) is made after 4 months after measurement and confirmation of performance (long time lag)
     • If the domain of performance measure was mostly out of clinicians' control
     • If the performance measure (payment scale) is a relative measure
   - Criteria for judging risk low:

     If the P4P program has 2 or more of the following features
     • If incentive payment (or penalty) is made before or at 4 months after measurement and confirmation of performance (short time lag)
     • If the domain of performance measure was mostly within the clinicians' control
     • If the performance measure (payment scale) is an absolute measure
   - Timing of payment after achieving targets (time lag): was it short or long?

     Short: If incentive payment (or penalty) is received not more than 4 months after measurement and confirmation of performance

     Long: If incentive payment (or penalty) is received more than 4 months after measurement and confirmation of performance
   - Domain of performance measured was the domain of performance measured within clinicians' control or out of clinicians' control?

     Criteria for judging within clinician's control: If incentive payments to health service providers are mostly/only based on processes and structures e.g. number of children immunized, routine measurement of blood pressure of patients every month, number of referrals made, rate of cancer screening

     Criteria for judging out of clinicians control: If incentive payments to health service providers depend on achieving a change in health outcomes e.g. reduction in mortality rates from a specific disease, blood pressure reduction, patient experience etc.
     Performance measure (payment scale) Absolute or relative measure?

     Absolute: If incentive is paid (fine levied) to the health service provider that based on their performance, not relative to how other health providers perform.

     Relative: If incentive payment is based on the performance of health service providers, relative to that of other providers.
9. Outcome /Impact