

## Access this article online

Quick Response Code:



Website:  
www.jehp.net

DOI:  
10.4103/jehp.jehp\_873\_23

# Evidence on the effectiveness of value-based payment schemes implemented in a hospital setting: A systematic review

Shahriyar Mokhtary, Ali Janati<sup>1</sup>, Mahmood Yousefi<sup>2</sup>, Behzad Raei<sup>3</sup>

Health Service Management, School of Management and Medical Informatics, Health Economics Department, Tabriz University of Medical Science, Tabriz, Iran, <sup>1</sup>Department of Health Policy and Management, Iranian Center of Excellence in Health Management, School of Management and Medical Informatics, Tabriz University of Medical Sciences, Tabriz, Iran, <sup>2</sup>Iranian Center of Excellence in Health Management, School of Management and Medical Informatics, Health Economics Department, Tabriz University of Medical Sciences, Tabriz, Iran, <sup>3</sup>Razi Educational and Therapeutic Center, Tabriz University of Medical Science, Tabriz, Iran

**Address for correspondence:**

Prof. Ali Janati,  
Department of Health Policy and Management, Iranian Center of Excellence in Health Management, School of Management and Medical Informatics, Tabriz University of Medical Sciences, Tabriz, Iran.  
E-mail: janati1382@gmail.com

Received: 20-06-2023  
Accepted: 02-09-2023  
Published: 29-08-2024

## Abstract:

Value-based payment is among payment models rewarding health care providers for achieving pre-defined targets of quality or efficiency measures of care. This paper aims to identify the evidence of the effectiveness of value-based payment schemes implemented in hospital settings. A systematic review of databases for studies published from 2000 to 2022 that evaluated VBP programs was conducted. We searched four databases including PubMed, Scopus, Embase, and Web of Sciences in July 2023. Studies were screened and assessed for eligibility. A thematic analysis approach was used to synthesize and summarize the findings. Overall, 29 articles looking into the VBP programs have been included. Most articles describe the effects on the outcome of care ( $n = 18$ ). The findings of a great deal of evidence in this field show that VBP is not correlated with some outcome measures including hospital-acquired conditions, 30-day mortality, mortality trends, as well as mortality among patients with acute myocardial infarction or heart failure. Only three of 12 studies have revealed a positive relationship between a P4P program and efficiency. Seven studies from the United States (US) found no evidence or mixed findings on the effects of P4P on efficiency. The magnitude of the effects of VBP on healthcare quality, patient experience, and costs has often been small and non-significant. The unintended negative impact of incentives in value-based payment on hospitals should be tackled when adopting policies and decisions.

## Keywords:

Efficiency, patient outcome assessment, quality, VBP

## Introduction

It has become apparent that healthcare systems are performing poorly in most countries. Resources used in inefficient ways have resulted in failures of needed care delivery. Moreover, the cost of health services is on the rise considerably. Governments, health insurance companies, and private sectors are unable to afford the provision of basic or essential health services. This problem becomes especially important when out-of-pocket payments increase.<sup>[1]</sup> Attempts are made to search for payment systems, which stimulate

offering well-organized, cost-effective, and high-quality care.

Financing hospitals or compensating clinical professionals through the predominant fee-for-service and diagnosis-related group (DRG) models underpinning volume are, therefore, generally thought to be an important reason for escalating costs of health systems.<sup>[2]</sup> Health policymakers are exploring new ways of paying for health improvement and enhancing patient value. Poor performance has encouraged initiative and new ideas that aim to bring efficiency, better outcomes, and positive patient experiences to health care delivery.<sup>[3]</sup>

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: WKHLRPMedknow\_reprints@wolterskluwer.com

**How to cite this article:** Mokhtary S, Janati A, Yousefi M, Raei B. Evidence on the effectiveness of value-based payment schemes implemented in a hospital setting: A systematic review. J Edu Health Promot 2024;13:327.

Starting in this century, pay-for-performance (P4P) models as alternative approaches have been introduced as a response. In P4P models, providers are reimbursed based on the accomplishment of given quality targets, which are mainly defined on the basis of process and structure indicators.<sup>[4]</sup>

A variant of P4P models is termed a value-based payment (VBP) program, encouraged mainly to be based on the outcome indicators, offered by the Centers for Medicaid and Medicare Services (CMS). Value is defined by achievement and improvement points for quality services subject to lower cost or price.<sup>[5]</sup> The VBP program was ratified in October 2012 as part of CMS's mission to improve healthcare quality. VBP is a payment mechanism for care that rewards hospitals and other providers for reaching quality standards or providing value in healthcare provision. Higher pay motive seems to modify the behavior of healthcare professionals.<sup>[6]</sup> Medicare hospitals are reimbursed by the indicator of VBP total performance score (TPS).<sup>[7]</sup> Under the VBP program, hospitals performing well on certain quality metrics are rewarded and those with poor performance are penalized. In 2013, VBP led to penalties for 1,427 hospitals and bonuses for 1,557 hospitals; embodying a redistribution of about \$1 billion among hospitals.<sup>[8]</sup>

Over 2013, two domains including clinical process of care and patient experience of care were value indicators.<sup>[9]</sup> Under the Inpatient Prospective Payment System (IPPS), Medicare payments were redistributed through the VBP program from poor-performing hospitals to better-performing hospitals. The percentage of Medicare reimbursements varied from 1% in 2013, 1.25% in 2014, 1.5% in 2015, and 1.75% in 2016 to 2% in 2017 and later years. However, it is important to make sure the funds used to improve financial performance may have been redirected from principal activities targeting community needs, such as programs for local health promotion, programs for the homeless, or the uninsured. This observation is particularly important when value-based purchasing becomes fairly widespread and, thus, the financial stakes for failing to meet performance targets become higher.<sup>[10]</sup> There are also concerns that performance score differentials may partly arise from variations in their patients' clinical or social characteristics, rather than only differences in the quality of care. Because budget neutrality provisions in these programs need bonuses and penalties to offset, inadequate risk adjustment could lead to unwarranted and sustained shifts of resources from procedures for sicker or more socially disadvantaged and excluded patients to procedures for healthier or a group of well-off patients.<sup>[11]</sup>

There are a few studies systematically assessing the VBP-related indicators. As far as we know, this was

the first study to assess specifically the VBP-related measures and the program effects as a new payment system. Previous research findings in this area have been inconsistent and contradictory. As a result, skepticism about VBP's effects on hospital performance creates a need for careful scrutiny of VBP's impacts. It will be critical to ensure that payment is structured in ways that actually drive improvement.<sup>[12]</sup> The aim of this review is to shine new light on these debates through an examination of the relationship between the VBP program and the predetermined measures in the relevant literature.

## Materials and Methods

A systematic review was conducted as a meta-analysis was not feasible due to the considerable heterogeneity among study outcome measures. We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.

### Databases and search strategy

A systematic search was conducted with four scientific databases, namely, Scopus, PubMed, EMBASE, and Web of Science. Databases were searched from their inception to July 2023. We first developed a list of search terms in consultation with a research librarian and used combinations of the following search terms and their synonyms:

"Health Care Category," "pay for quality," "pay for value," "value-based payment," "pay for outcome," "pay for output," "value-based purchasing," "value-based reimbursement," "fee for value," and "pay for performance." Search strategies were designed to be suitable to the specific features of each database. In addition, the reference lists of articles that were selected via the search strategy were hand-searched to include any further articles that may have been missed.

### Inclusion and exclusion criteria

Studies were included if they were (1) written in the English language, (2) from a peer-reviewed academic journal, (3) conducted in a hospital setting, and (4) contained the results of an effect of VBP schemes on organizational and patient outcomes. Reviews, letters, commentaries, dissertations, and non-scholarly sources were also excluded from the review. In addition, articles that did not examine the effects of VBP schemes were excluded. For example, our search strategy in PubMed was as follows:

("Health Care Category"[Mesh]) AND (("Value-based payment"[Title/Abstract] OR "pay for quality"[Title/Abstract] OR "pay for value"[Title/Abstract] OR "pay for outcome"[Title/Abstract] OR "pay for output"[Title/

Abstract] OR “value-based purchasing”[Title/Abstract] OR “value-based reimbursement”[Title/Abstract] OR “fee for value”[Title/Abstract] OR “pay for performance”[Title/Abstract] OR “P4P”)[Title/Abstract].

### Data extraction and quality assessment

Studies were included if they examined the effectiveness of VBPs on any type of outcome (e.g., clinical process, patient outcomes, expenditures, quality of care, efficiency, utilization, patient experience, and patient safety). Data items from each eligible study were abstracted, and reviewed in terms of accuracy by one additional investigator. We extracted the information on study design and setting, sample size, country, main objectives, type of statistical analysis, and sample size, as well as the association between the incentive program and the determined measures. Two reviewers independently assessed study quality using the Newcastle–Ottawa Scale, a generic and widely applied method developed by Dow- and Black.<sup>[13]</sup> Disagreements were resolved in a discussion with a third reviewer. In the Dow- and Black method, articles receive points on 27 items covering four domains: reporting, external validity, internal validity, and power. The more points an article receives, the higher the methodological quality of the article. The maximum number of points was 32. The quality assessment score in this review for included studies ranged between 8 and 17. We excluded articles with a quality score of less than 8. We employed this generic appraisal tool due to higher heterogeneity in the statistical designs of included studies (e.g., before-after design, cross-sectional studies, and interrupted time series).<sup>[14]</sup>

### Data analysis

After the data extraction, extracted data were categorized and analyzed. A directed qualitative content analysis was used for data analysis. This directed qualitative content analysis was based on the Hsieh and Shannon approach. We extracted data on the VBP models studied and the outcomes of these models. Outcomes were divided into following four themes: clinical outcomes, patient-reported outcomes/experiences, organization-related outcomes/experiences, and cost outcomes.

## Result

The database search retrieved 8,875 records. After removing duplicates, 5,093 titles and abstracts were screened, and 783 records were reviewed in full text. We selected 52 relevant studies for assessing quality. After quality assessment, 29 articles were included in the final review [see Figure 1. Flow of information through the review -PRISMA chart].

### Characteristics of studies

Included studies were published in 2008 and comprised a variety of designs: cross-sectional analysis ( $n = 5$ ), longitudinal panel data analyses ( $n = 4$ ), an interrupted time series ( $n = 5$ ), and a retrospective cohort ( $n = 1$ ), a retrospective, quasi-experimental design ( $n = 1$ ), and pre-post study design with a comparison two groups ( $n = 3$ ). Most included studies were from North America including the US ( $n = 22$ ; 75.9%) and three studies originated from Africa including Tanzania ( $n = 2$ ) and Zimbabwe ( $n = 1$ ). Three studies were from Europe, with one study each originating from Italy, England, and France. One study was from Asia including Lebanon.

This work is to review recent research into the effects of value-based payments; as a result, we considered some measures to assess the possible benefits of VBP. The results obtained from the reviewed evidence are presented in Table 1. Hospital settings or groups of hospitals with distinguishing characteristics were considered as a unit of our analysis.

- Premier Hospital Quality Incentive Demonstration (PHQID) program
- Value-based purchasing program
- Non-payment for hospital-acquired conditions program
- Hospital readmissions reduction program (HRRP)
- Advancing quality program
- Financial incentive to quality improvement program (FIQIP).

Characteristics of the reviewed studies and the relationship between HVBP and outcomes have been represented in Table 2. The main outcomes were defined into five categories as follows:

### Clinical outcomes

Because included studies differ considerably in terms of examined outcomes, we report the results for each separately. Four controlled pre-post studies under the value-based purchasing program were included in this review. The findings of these studies showed that improvements in the clinical process of care as a prominent provision measure were small and not significantly greater among hospitals under VBP than among control groups.<sup>[15,16]</sup> However, one study found that the program was significantly associated with improved performance for the two clinical process measures related to pneumonia.<sup>[15]</sup> Spaulding (2014) found that VBP does not appear to correlate with improved quality.<sup>[6]</sup> One controlled before–after study evaluated the impact of California DSRIP’s impact on central line-associated bloodstream infections, venous thromboembolism, and hospital-acquired pressure ulcers. It found no impact.<sup>[17]</sup> Provision outcomes underpay for performance programs in the form of

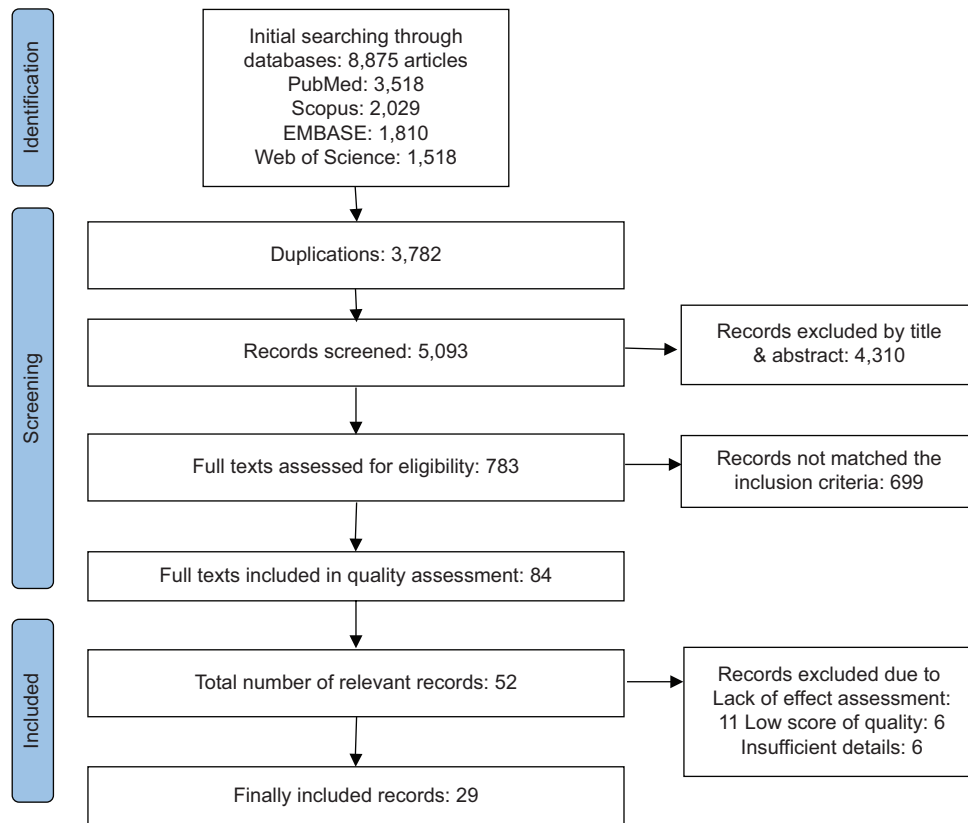


Figure 1: Flow of information through the review

various quality measures have been reported in some studies. Fichera (2021) evaluated a results-based financing program in Zimbabwe on health outcomes and found no evidence of an effect on any other incentivized indicators such as four antenatal care visits, full immunization or TT + vaccination, HIV testing, family planning, postnatal care, and vitamin A delivery.<sup>[18]</sup> Lalloué (2017) evaluated the effects of financial incentives to quality improvement (IFAQ) on quality indicators including the quality and content of the medical record, and screening for nutritional disorders, in France and reported that the difference-in-differences effect was positive but not significant in the crude model. In this study, crude models showed no significant effects, and only adjusted models detected a small effect on pain assessment traceability.<sup>[19]</sup> The effects of bonus and penalty sizes on provision outcomes in P4P programs have been assessed in some studies. In one study, Lee (2019) found that penalized hospitals tend to improve both clinical processes and outcomes.<sup>[3]</sup> Likewise, the other study revealed that double bonuses in the Medicare Advantage Quality Bonus Payment Demonstration were not associated with improved quality.<sup>[20]</sup> Mellor (2016) studied the effect of HRRP under three conditions and found that the program significantly reduced the likelihood of readmission for Medicare patients treated for AMI; however, PN and HF readmissions are unaffected by the HRRP.<sup>[21]</sup>

### Patient-reported outcomes/experience

In the literature, patient outcome refers to health service utilization and resulting changes in patient health status such as mortality, morbidity, satisfaction, and some adverse clinical events that are beyond the provider's full control.<sup>[22]</sup> The most commonly reported measures capture patient experience and mortality. Patient experience assessment among other patient outcomes addresses the dimensions of the area of communication with physicians, the responsiveness of hospital personnel, pain management, hospital environment conditions in terms of cleanliness and quietness, and after-discharge satisfaction.<sup>[3]</sup> Among the 29 studies included in this review, 16 studies reported patient outcomes. Three articles have examined the effect of VBP on patient experience, one study with difference-in-difference analysis,<sup>[15]</sup> and two studies with interrupted time-series design.<sup>[16,23]</sup> Their main findings showed that VBP was not associated with improved patient experience measures. One study evaluated the impact of VBP on patient safety and indicated that the program did not appear to correlate with improved patient safety.<sup>[6]</sup> Binyaruka (2015) evaluated the impact of P4P in Tanzania and found no evidence of an effect of P4P on patient experience of care for targeted services.<sup>[24]</sup> In this regard, Lee (2019) reported that more heavily penalized hospitals focus on perceived patient satisfaction.<sup>[3]</sup> Ten out of 31 articles in the current review

**Table 1: Characteristics of included studies**

Author/year	Country	Design of study	Setting/sample	Type of payment Program	Targeted outcomes	QS'
Hsu <i>et al.</i> , 2020	USA	An interrupted time series (Hospital value-based purchasing (HVBP) implementation in fiscal year 2015 or 2016)	618 hospitals (145 safety-net vs 473 non-safety-net hospitals)	1-HVBP	Health care-associated infections	15 (1)
Lee <i>et al.</i> , 2019	USA	Longitudinal panel data analyses (Period 2013–2016) (Two-stage least squares)	2,861 VBP participating hospitals	HVBP	1-Performance score, 2-Patient experiential quality, 3-Clinical performance measure, 4-Change in Case Mix Index	11.5 (2.5)
Bazzoli <i>et al.</i> , 2018	USA	Pooled cross-sectional data	4,824 hospital-year observations	Hospital Readmission Reduction Program and HVBP	1-Financial measures, 2- Operating margin, 3- Total margins	11 (0)
Papanicolas <i>et al.</i> , 2017	USA	Interrupted time series	3,452 hospitals (419 non-VBP + 3033 VBP Hospitals)	HVBP	1- Patient Experience	10 (1)
Ryan <i>et al.</i> , 2015	USA	Difference-in-differences analysis (Longitudinal data, from 2011 through 2012)	Clinical process performance (2,801 VBP hospitals + 240 comparison hospitals) patient experience performance (2,779 VBP hospitals + 284 comparison hospitals)	HVBP	1-Clinical process (12 Measures) 2-Patient experience (8 Measures)	12 (0)
Spaulding <i>et al.</i> , 2014	USA	Zero-inflated negative binomial regression	2927 hospitals	HVBP	1- Quality outcome and patient safety	15.5 (0.5)
Roberts <i>et al.</i> , 2019	USA	Observational design (a cross-sectional regression discontinuity design)	practices with $\geq 100$ clinicians ( $n=931$ ) practices with $\geq 10$ clinicians ( $n=8,491$ )	HVBP	1-Hospitalization for ambulatory care-sensitive conditions 2- all-cause 30-day readmissions, 3- Medicare spending 4- Annual mortality	11 (1)
Lalloué <i>et al.</i> , 2017	France	Cross-sectional analysis (difference-in-differences method)	- Intervention group ( $n=185$ Hospitals) - Control group ( $n=192$ Hospitals)	the P4P program IFAQ	Quality indicators	11 (1)
Karim <i>et al.</i> , 2020	USA	Longitudinal unbalanced panel dataset from 2008 to 2015.	24,517 hospital-year observations	1) HVBP 2) HRRP	1-Operating margin and 2- total margin	10 (1)
Chen <i>et al.</i> , 2017	USA	7-year panel dataset Pre-period (2008–2010) post-period 1 (2011–2012) post-period 2 (2013– 2014)	1) Delta Hospitals ( $n=1274$ ) 2) Non-Delta Hospitals ( $n=19,553$ )	1) HVBP 2) HRRP	1-operating margin, 2- total margin	13 (1)
Layton <i>et al.</i> , 2015	USA	Difference-in-differences analysis (Treatment group, Control group, Matched Control)	(Treatment group ( $n=197$ ), Control group ( $n=1125$ ), Matched Control (532))	P4P	1-Quality of Care and 2-Number of Plans Offered among Counties	10 (0)
Calikoglu <i>et al.</i> , 2015	USA	An interrupted time series (2007-2010)	forty-six hospitals	Quality-Based Reimbursement Program, Hospital-Acquired Conditions Program	Clinical process- of-care	10 (0)
Bastian <i>et al.</i> , 2016	USA	A retrospective, quasi-experimental design (during the period of 2001–2012.)	23 Army hospitals, 12 Air Force hospitals, and 19 Navy hospitals,	P4P	Technical efficiency	10 (0)
Kristensen <i>et al.</i> , 2008	England	Cross-sectional (difference-in-differences regression analysis)	161 hospitals (24 hospitals in the northwest region and 137 elsewhere in England)	P4P	30-day in-hospital mortality	11 (1)

Contd...

**Table 1: Contd...**

Author/year	Country	Design of study	Setting/sample	Type of payment Program	Targeted outcomes	QS <sup>1</sup>
Cox <i>et al.</i> , 2015	USA	Experimental study	1 hospital (30 representative patients)	P4P	1-Hospital readmission rates, 2-Hospital length of stay)	14 (2)
Binyaruka <i>et al.</i> , 2020	Tanzania	Two rounds of a repeated cross-sectional facility survey	75 facilities (P4P arm) and 75 facilities (comparison arm)	P4P	Technical efficiency	15 (2)
Khalife <i>et al.</i> , 2016	Lebanon	A single-group interrupted time series analysis model	146 public and private hospitals	P4P	1-Medical case mix, 2-Surgical case mix, and 3-Mixed cases	17 (1)
Fichera <i>et al.</i> , 2021	Zimbabwe	Difference in differences design	1533 health care facilities (across 62 districts)	Result Based Financing	Health outcomes	15 (1)
Glickman <i>et al.</i> , 2009	USA	Observational study (principal components analysis, regression analyses)	4226 hospitals	PHQID	Process measures	13 (0)
Banerjee <i>et al.</i> , 2019	USA	Difference-in-differences models	1902 hospitals in each year between 2009 and 2016	HVBP	30-day mortality	16 (1)
Izón <i>et al.</i> , 2018	USA	Observational study (Translog-specification of a stochastic cost frontier)	290 California hospitals for 2012–2015	HVBP	1-Performance score, 2-Mortality rate, 3-Readmission rate, 4-Cost efficiency	15 (0)
Figueroa <i>et al.</i> , 2016	USA	Observational study (Difference in Difference, longitudinal data from 2008 to 2013)	4267 acute care hospitals (2919 HVBP hospitals versus 1348 non-HVBP hospitals)	HVBP	30 day mortality	16 (1)
Ryan <i>et al.</i> , 2017	USA	An interrupted time series 2008 to 2015	Exposed hospitals (1364 to 2615) and -control hospitals (31 to 617)	HVBP	1-Clinical process and 2-patient experience, 3-30-day risk-standardized mortality	13 (2)
Peluso <i>et al.</i> , 2019	Italy	Multivariate difference in difference approach	142 hospitals during the four years 2010–2013	P4P	Health outcomes	12 (0)
Mellor <i>et al.</i> , 2016	USA	The triple difference design	67 hospitals in the AMI analysis and 71 hospitals in the HF and PN analyses.	HRRP	Readmission	8.5 (0.5)
Ryan <i>et al.</i> , 2014	USA	Regression discontinuity design (from 2004 to 2006)	260 hospitals participating in the PHQID	PHQID	Composite quality measures the level of quality in acute myocardial infarction, pneumonia, and heart failure	10 (0)
Binyaruka <i>et al.</i> , 2015	Tanzania	A controlled before and after study design (Difference in difference)	150 facilities (Intervention group) and 75 facilities Comparison group	P4P program	1-Service Utilization, 2-paying out of pocket for deliveries, 3-the provision of anti-malarial during pregnancy, 4- Patient experience, and patient satisfaction	15 (2)
Ryan, 2009	USA	A retrospective cohort of 11,232,452 admissions from 6,713,928 patients	3,570 acute care hospitals between 2000 and 2006	PHQID	1-30-day mortality and 2-60-day cost	12 (1)
Keller <i>et al.</i> , 2021	USA	Pre-post study design with a comparison two group (Designated Public Hospitals (DPHs) hospitals and nonDPHs)	2009-2014 discharge data from California hospitals	DSRIP	adverse outcomes	14 (1)

<sup>1</sup>Quality assessment

have reported the effects of different programs on mortality. The effect of the VBP on mortality measures has been addressed in five studies.<sup>[11,16,25-27]</sup> The findings

of a great deal of evidence in this field show that VBP is not correlated with a long-term and significant effect on mortality. However, Ryan *et al.*,<sup>[16]</sup> demonstrated

that VBP was associated with a significant reduction in mortality among patients who were admitted for pneumonia. Other studies have considered the relationship between P4P programs and mortality. They did not find significant changes in mortality that can be attributed to the P4P adoption.<sup>[17,28-30]</sup> One study by Fichera in Zimbabwe found that result-based financing programs reduced under-five mortality by two percentage points.<sup>[18]</sup> One study conducted in Italy showed that two out of the five health outcomes considered, that is, readmissions and transfers, support the hypothesis that the P4P introduction had a positive effect on quality.

### Cost outcomes

Hospital costs account for a large portion of the total healthcare expenditures. Payment methods for hospitals can affect the quality and efficiency of service provision. Healthcare providers can potentially be incentivized to deliver care to maximize patient benefit while managing and controlling resource use.<sup>[31]</sup> In the past, policy interventions in reimbursement were based on cost containment;<sup>[32]</sup> however, new payment methods are designed with the aim of improving performance (efficiency and quality).<sup>[33]</sup> One study found negative and highly significant associations between the VBP penalty rate and both operating and total margins,<sup>[10]</sup> whereas a study by Roberts (2019) showed no significant relationship between financial incentives and Medicare spending.<sup>[11]</sup> Another study found that the relationships between Magnet recognition, outcome, and efficiency were not statistically significant.<sup>[7]</sup> A study by Izón (2017)<sup>[34]</sup> reaches different conclusion, finding an association between quality scores and increased operating costs. A retrospective, quasi-experimental design showed a statistically significant reduction in technical efficiency for the hospitals that received pay-for-performance financial incentives.<sup>[35]</sup> Nonetheless, Cox, et al.<sup>[36]</sup> suggest that P4P mechanisms incentivize cost-effective reductions in hospital readmission rates. In another major study, Binyaruka and Anselmi, (2020)<sup>[37]</sup> reported that there is no evidence of a P4P effect on efficiency on average. According to Ryan (2009)<sup>[30]</sup>, the HQID program had no effect on the risk-adjusted 60-day cost for acute myocardial infarction, heart failure, pneumonia, or coronary-artery bypass grafting. The evidence from a study suggests P4P was associated with a 5.0% reduction in those paying out-of-pocket for deliveries; however, no evidence of an effect on the average amount paid was found.<sup>[24]</sup> Two studies examining the financial effect of value-based purchasing and hospital readmission reduction programs reached similar conclusions. The findings from these studies suggest that there is a significant decrease in operating margins for some hospitals after the implementation of HRRP and VBP.<sup>[38,39]</sup> Based on these results, most

authors conclude that P4P effects on resource use have been inconsistent, and convincing evidence is required.

### Organization-related outcomes/experience

No relevant healthcare provider outcomes were reported. Only one study has reported that the bundled payment mechanism has led to the largest reduction in patient length of stay without changing the probability of readmission.<sup>[36]</sup>

## Discussion

This systematic review of 29 studies providing valuable evidence about the effectiveness of VBP programs is an update on previous reviews. Overall, in the hospital setting, the current review found that the impact on the improvement of process-of-care outcomes was controversial among the studies. These studies were conducted in the US, where baseline performance is rich and health systems have little incentive. These mixed results fail to provide conclusive evidence to support the effectiveness of using financial incentives to drive the clinical process. One study on the effects of the hospital-acquired conditions program found some evidence that the program improved performance on clinical processes including heart failure and acute myocardial infarction.<sup>[41]</sup> The result of a study showed no evidence that the HVBP improved the clinical process, although there was a significant association between VBP and improved performance for the two clinical process measures related to pneumonia.<sup>[15]</sup> It is noteworthy that some leading experts recommend that process measures are of great significance for assessment because they are under the hospital's control and do not necessarily lead to improved patient outcomes.<sup>[44]</sup>

The impact of VBP on the patient experience was reported in some studies. Although some research has found positive effects in relation to VBP programs, the literature has emerged that offers contradictory findings across studies. These rather contradictory results may be due to the fact that the observational nature of studies makes it hard to attribute some changes methodologically to the implementation of P4P programs. For instance, studies with before-after designs have a higher proportion of statistically significant effects compared to the studies with interrupted time series and differences in different designs.<sup>[45]</sup> It also has been maintained that the limited impact of P4P programs has come largely from the small size of the financial incentive.<sup>[46]</sup> The major flaw encountered in patient experience indicators is that they might misrepresent the actual quality of care because of the subject nature of patients' attitudes, which appear to differ by demographic characteristics.<sup>[27]</sup> In accordance with the present results, a previous review has demonstrated that P4P programs generally did not

**Table 2: Findings from studies of relationships between VBP and selected measures**

Study	Measures			
	Patient-reported outcomes/experience	Organization-related outcomes/experience	Clinical outcomes	Cost outcomes
E. Hsu <i>et al.</i> , 2020 <sup>[40]</sup>	-	-	No effect/association	-
Lee <i>et al.</i> , 2020 <sup>[3]</sup>	Positive	Positive	Positive	Positive
Bazzoli <i>et al.</i> , 2018 <sup>[10]</sup>	-	-	-	Mixed effect/association
Papanicolas <i>et al.</i> , 2017 <sup>[23]</sup>	Positive	-	-	-
Ryan <i>et al.</i> , 2015 <sup>[15]</sup>	No effect/association	-	No effect/association	-
Spaulding <i>et al.</i> , 2014 <sup>[6]</sup>	No effect/association	-	No effect/association	-
Roberts, 2018 <sup>[11]</sup>	No effect/association	-	-	No effect/association
Lalloué <i>et al.</i> , 2017 <sup>[19]</sup>	-	No effect/association	No effect/association	-
Karim <i>et al.</i> , 2021 <sup>[39]</sup>	-	-	-	Mixed effect/association
Chen <i>et al.</i> , 2017 <sup>[38]</sup>	-	-	-	Mixed effect/association
Layton <i>et al.</i> , 2015 <sup>[20]</sup>	-	Mixed effect/association	-	-
Calikoglu <i>et al.</i> , 2012 <sup>[41]</sup>	Positive	-	-	Positive
Bastian <i>et al.</i> , 2016 <sup>[35]</sup>	-	Negative	-	-
Kristensen <i>et al.</i> , 2014 <sup>[29]</sup>	Mixed effect/association	-	-	-
Cox <i>et al.</i> , 2016 <sup>[36]</sup>	-	No effect/association	Positive	-
Binyaruka <i>et al.</i> , 2021 <sup>[37]</sup>	-	No effect/association	-	-
Khalife <i>et al.</i> , 2020 <sup>[42]</sup>	-	Positive	-	-
Fichera <i>et al.</i> , 2021 <sup>[18]</sup>	Positive	-	-	-
Glickman <i>et al.</i> , 2009 <sup>[27]</sup>	Positive	-	-	-
Banerjee <i>et al.</i> , 2019 <sup>[25]</sup>	No effect/association	-	-	-
Izón <i>et al.</i> , 2017 <sup>[34]</sup>	Positive	-	-	Positive
Figueroa <i>et al.</i> , 2016 <sup>[26]</sup>	No effect/association	-	-	-
Ryan <i>et al.</i> , 2017 <sup>[16]</sup>	No effect/association	-	-	-
Peluso <i>et al.</i> , 2019 <sup>[28]</sup>	Mixed	-	-	-
Mellor <i>et al.</i> , 2017 <sup>[21]</sup>	Positive	-	-	-
Ryan <i>et al.</i> , 2014 <sup>[43]</sup>	-	-	No effect/association	-
Binyaruka <i>et al.</i> , 2015 <sup>[24]</sup>	No effect/association	Positive	Positive	No effect/association
Ryan, 2009 <sup>[30]</sup>	No effect/association	-	-	No effect/association
Keller <i>et al.</i> , 2021 <sup>[17]</sup>	-	-	No effect/association	-

result in an improved patient experience or immediate changes in patient experience in hospital settings.<sup>[47]</sup> In the analysis of the P4P effect, Ryan *et al.*<sup>[16]</sup> (2017) concluded that patient experience measures have reduced a bit following the adoption of the HVBP program.

The effect of P4P programs on patient outcomes has been reported in most studies. Ten studies have assessed patient outcome indicators under HVBP. Despite the inconsistent results, we found little to no effect on patient outcome metrics. Studies show striking results, suggesting HVBP implementation has not been causally associated with reduced infections,<sup>[40]</sup> lower mortality,<sup>[25,26]</sup> and improved quality as a main outcome.<sup>[6]</sup> One important study also showed that VBP was not associated with significantly reduced mortality among patients admitted for acute myocardial infarction and heart failure, although it was associated with reduced mortalities from pneumonia.<sup>[16]</sup> The improvements in patient outcome indicators have been reported under other P4P programs in a number of studies. However, the reviewed literature found little evidence for the positive effect on outcomes. The magnitude of the effect has often

been small and non-significant, and it was demanding to attribute the size of the changes in outcomes to a given intervention due to the presence of time-varying confounders and methodological flaws of some observational studies in isolating the effects. It is difficult to explain this result; however, it might be related to the lack of adequate financial incentives under the HVBP. Another possible explanation for this is that studies fail to assess long-term outcomes, requiring lasting follow-up data.<sup>[25]</sup> In addition, P4P programs may be implemented alongside other quality improvement interventions such as electronic decision-support and audit feedback, thus making it complicated to differentiate the incremental effect of P4P.<sup>[47]</sup>

In accordance with these results, previous reviews have demonstrated that P4P programs, including HVBP, do not lead to substantial improvements in patient outcomes or the quality of care.<sup>[47,48]</sup> Literature on P4P programs shows results that agree with our findings on patient outcomes. Findings from an overview study based on a review of four systematic reviews show that financial incentives have generally a positive



effect on the process of care; however, there was a lack of evidence of improvement in patient outcomes.<sup>[49]</sup> Nonetheless, much uncertainty still exists about the relationship between value-based payment effects and patient outcomes including mortality or quality of care. Moreover, improvements are temporary and emerge mainly over the first 3 years of implementation and then it will be attenuated by the quality of care for non-incentivized conditions, patient-centered care, and suffered continuity of care.<sup>[50]</sup> Some studies have demonstrated that the magnitude of bonuses or penalties in P4P schemes or hospital characteristics were effective on targeted measures,<sup>[3,15]</sup> whereas, robust evidence from an inclusive review study found no key effect modifier for the effectiveness of P4P programs.<sup>[51]</sup>

The findings in this review are subject to a number of limitations. First, the current systematic review was not specifically designed to evaluate factors related to the success of P4P programs such as the type or size of hospitals, region, and design of VBP schemes. Second, it is limited by the lack of information on the time period from the onset of the program to the assessment of measures. Third, we included some cross-sectional designs without having comparison groups and assessing risk of bias in them. Some characteristics that are relevant to patient outcomes necessitate being cautious in making any inferences about the results. In addition, relying on a given medical condition may lead to bias in terms of the generalization of the results to some other conditions. There is abundant room for further progress in determining how to improve the relationship between VBP and related measures, as well as determine how we can link evidence-based medical practices to the VBP system.

## Conclusion

The present study was designed to determine the effect of value-based payment programs in hospital settings to better inform policymakers and opinion leaders to distill policy recommendations. This research confirms the previous findings and contributes additional evidence that suggests these programs have generated a lot of research activity and many pundits' attention to address the concerns of value-based payments; however, the system might not see the slightest improvement. In general, therefore, it seems that most of these programs are fads. It is also worth noting that the effect of payment methods might be sensitive to context and setting elements, although the mechanisms through which they exert their effects are complex and not easily measurable. In other words, isolating the effect of an adoption of a new payment model alongside other programs such as hospital monitoring, where confounding is a major concern, can be a challenge.

The main issue in undertaking payment interventions may be to satisfy some conditions, under which costs and consequences of the program are balanced and at the same time the interests of a range of stakeholders are fulfilled.

## List abbreviations

VBP: value-based payment, DRG: diagnosis-related group, P4P: pay-for-performance, CMS: Centers for Medicaid and Medicare Services, TPS: total performance score, IPPS: inpatient prospective payment system, PHQID: premier hospital quality incentive demonstration, HRRP: hospital readmissions reduction program, FIQIP: financial incentive to quality improvement program.

## Authors' contributions

SM and AJ conceptualized the study. SM and BR contributed to database search and formal analysis. Project administration and supervision was performed by AJ and MY together, and SM was a major contributor to writing the manuscript. All authors read and approved the final manuscript.

## Acknowledgments

The authors would like to gratefully acknowledge the assistance of all experts who made comments and helpful suggestions regarding the conduct of this study.

## Ethical consideration

The present study is part of the first author's doctoral thesis in health service management at the Tabriz University of Medical Sciences that was conducted at and confirmed by the Research Ethics Committees of the Tabriz University of Medical Sciences with approval ID: IR.TBZMED.REC.1401.358.

## Financial support and sponsorship

This research was supported by the Tabriz University of Medical Sciences.

## Conflicts of interest

There are no conflicts of interest.

## References

1. 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. *Social Science & Medicine*. 2021;270:113624.
2. Orszag PR, Ellis P. The challenge of rising health care costs—a view from the Congressional Budget Office. *New Engl J Med* 2007;357:1793-5.
3. Lee SJ, Venkataraman S, Heim GR, Roth AV, Chilingirian J. Impact of the value-based purchasing program on hospital operations outcomes: An econometric analysis. *J Oper Manag* 2020;66:151-75.

4. Nicholson S, Pauly MV, Wu AYJ, Murray JF, Teutsch SM, Berger ML. Getting real performance out of pay-for-performance. *Milbank Q* 2008;86:435-57.
5. Shortell SM, Kaluzny AD. Organization theory and health services management. *Essentials of Health Care Management*. Thompson Delmar Learning, New York 1994. p. 17-43.
6. Spaulding A, Zhao M, Haley DR. Value-based purchasing and hospital acquired conditions: Are we seeing improvement? *Health Policy* 2014;118:413-21. doi.org/10.1016/j.socscimed.2020.113624.
7. Lasater KB, Germack HD, Small DS, McHugh MD. Hospitals known for nursing excellence perform better on value based purchasing measures. *Policy Politics Nurs Pract* 2016;17:177-86.
8. Centers for Medicare and Medicaid Services (CMS), HHS. Medicare program; hospital inpatient prospective payment systems for acute care hospitals and the long-term care hospital prospective payment system and Fiscal Year 2014 rates; quality reporting requirements for specific providers; hospital conditions of participation; payment policies related to patient status. Final rules. *Fed Regist* 2013;78:50495-1040.
9. Centers for Medicare & Medicaid Services, "CMS issues final rule for first year of hospital value-based purchasing program.", (2011), available from <https://www.cms.gov/newsroom/fact-sheets/cms-issues-final-rule-first-year-hospital-value-based-purchasing-program>.
10. Bazzoli GJ, Thompson MP, Waters TM. Medicare payment penalties and safety net hospital profitability: Minimal impact on these vulnerable hospitals. *Health Serv Res* 2018;53:3495-506.
11. Roberts ET, Zaslavsky AM, McWilliams JM. The value-based payment modifier: Program outcomes and implications for disparities. *Ann Intern Med* 2018;168:255-65.
12. Rau J. Nearly 1,500 hospitals penalized under medicare program rating quality. *Kaiser Health News*. 2013.
13. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health* 1998;52:377-84.
14. Vlaanderen F, Tanke M, Bloem B, Faber M, Eijkenaar F, Schut F, et al. Design and effects of outcome-based payment models in healthcare: A systematic review. *Eur J Health Econ* 2019;20:217-32.
15. Ryan AM, Burgess JF Jr, Pesko MF, Borden WB, Dimick JB. The early effects of Medicare's mandatory hospital pay-for-performance program. *Health Serv Res* 2015;50:81-97.
16. Ryan AM, Krinsky S, Maurer KA, Dimick JB. Changes in hospital quality associated with hospital value-based purchasing. *N Engl J Med* 2017;376:2358-66.
17. Keller MS, Chen X, Godwin J, Needleman J, Pourat N. Evaluating inpatient adverse outcomes under California's delivery system reform incentive payment program. *Health Serv Res* 2021;56:36-48.
18. Fichera E, Anselmi L, Gwati G, Brown G, Kovacs R, Borghi J. Can results-based financing improve health outcomes in resource poor settings? Evidence from Zimbabwe. *Soc Sci Med* 2021;279:113959. doi: 10.1016/j.socscimed.2021.113959.
19. Lalloué B, Jiang S, Girault A, Ferrua M, Loirat P, Minvielle E. Evaluation of the effects of the French pay-for-performance program—IFAQ pilot study. *Int J Qual Health Care* 2017;29:833-7.
20. Layton TJ, Ryan AM. Higher incentive payments in Medicare advantage's pay-for-performance program did not improve quality but did increase plan offerings. *Health Serv Res* 2015;50:1810-28.
21. Mellor J, Daly M, Smith M. Does it pay to penalize hospitals for excess readmissions? Intended and unintended consequences of Medicare's hospital readmissions reductions program. *Health Econ* 2017;26:1037-51.
22. Yuan B, He L, Meng Q, Jia L. Payment methods for outpatient care facilities. *Cochrane Database Syst Rev* 2017;3:CD011153. doi: 10.1002/14651858.CD011153.pub2.
23. Papanicolas I, Figueroa JF, Orav EJ, Jha AK. Patient hospital experience improved modestly, but no evidence Medicare incentives promoted meaningful gains. *Health Aff* 2017;36:133-40.
24. Binyaruka P, Patouillard E, Powell-Jackson T, Greco G, Maestad O, Borghi J. Effect of paying for performance on utilisation, quality, and user costs of health services in Tanzania: A controlled before and after study. *PLoS One* 2015;10:e0135013. doi: 10.1371/journal.pone.0135013.
25. Banerjee S, McCormick D, Paasche-Orlow MK, Lin M-Y, Hanchate AD. Association between degree of exposure to the hospital value based purchasing program and 30-day mortality: Experience from the first four years of medicare's pay-for-performance program. *BMC Health Serv Res* 2019;19:1-12. doi: 10.1186/s12913-019-4562-7.
26. Figueroa JF, Tsugawa Y, Zheng J, Orav EJ, Jha AK. Association between the value-based purchasing pay for performance program and patient mortality in US hospitals: Observational study. *BMJ* 2016;353:i2214. doi: 10.1136/bmj.i2214.
27. Glickman SW, Boulding W, Roos JM, Staelin R, Peterson ED, Schulman KA. Alternative pay-for-performance scoring methods: Implications for quality improvement and patient outcomes. *Med Care* 2009;47:1062-8.
28. Peluso A, Berta P, Vinciotti V. Do pay-for-performance incentives lead to a better health outcome? *Empir Econ* 2019;56:2167-84.
29. 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.
30. Ryan AM. Effects of the premier hospital quality incentive demonstration on medicare patient mortality and cost. *Health Serv Res* 2009;44:821-42.
31. 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;110:115-30.
32. Dixon J. Payment by results—new financial flows in the NHS. *BMJ* 2004;328:969-70.
33. Kondo KK, Damberg CL, Mendelson A, Motu'apuaka M, Freeman M, O'Neil M, et al. Implementation processes and pay for performance in healthcare: A systematic review. *J Gen Intern Med* 2016;31:61-9.
34. Izón GM, Pardini CA. Association between Medicare's mandatory hospital value-based purchasing program and cost inefficiency. *Appl Health Econ Health Policy* 2018;16:79-90.
35. Bastian ND, Kang H, Griffin PM, Fulton LV. Measuring the effect of pay-for-performance financial incentives on hospital efficiency in the military health system. *IJSE Trans Healthc Syst Eng* 2016;6:33-41.
36. Cox JC, Sadiraj V, Schnier KE, Sweeney JF. Incentivizing cost-effective reductions in hospital readmission rates. *J Econ Behav Organ* 2016;131:24-35.
37. Binyaruka P, Anselmi L. Understanding efficiency and the effect of pay-for-performance across health facilities in Tanzania. *BMJ Glob Health* 2020;5:e002326. doi: 10.1136/bmjgh-2020-002326.
38. Chen H-F, Karim S, Wan F, Nevola A, Morris ME, Bird TM, et al. Financial performance of hospitals in the Mississippi delta region under the hospital readmissions reduction program and hospital value-based purchasing program. *Med Care* 2017;55:924-30.
39. Karim SA, Nevola A, Morris ME, Tilford JM, Chen HF. Financial performance of hospitals in the appalachian region under the hospital readmissions reduction program and hospital value-based purchasing program. *J Rural Health* 2021;37:296-307.
40. Hsu HE, Wang R, Broadwell C, Horan K, Jin R, Rhee C, et al. Association between federal value-based incentive programs and health care-associated infection rates in safety-net and non-safety-

- net hospitals. *JAMA Netw Open* 2020;3:e209700. doi: 10.1001/jamanetworkopen.2020.9700.
41. Calikoglu S, Murray R, Feeney D. Hospital pay-for-performance programs in Maryland produced strong results, including reduced hospital-acquired conditions. *Health Aff* 2012;31:2649-58.
  42. Khalife J, Ammar W, Emmelin M, El-Jardali F, Ekman B. Hospital performance and payment: Impact of integrating pay-for-performance on healthcare effectiveness in Lebanon. *Wellcome Open Res* 2020;5:95.
  43. Ryan A, Sutton M, Doran T. Does winning a pay-for-performance bonus improve subsequent quality performance? Evidence from the hospital quality incentive demonstration. *Health Serv Res* 2014;49:568-87.
  44. Nardin R, Zallman L, McCormick D, Woolhandler S, Himmelstein D. The uninsured after implementation of the affordable care act: A demographic and geographic analysis. *Health Aff* 2013. doi: 10.1377/forefront.20130606.031785.
  45. Zaresani A, Scott A. Is the evidence on the effectiveness of pay for performance schemes in healthcare changing? Evidence from a meta-regression analysis. *BMC Health Serv Res* 2021;21:1-10.
  46. Milstein R, Schreyoegg J. Pay for performance in the inpatient sector: A review of 34 P4P programs in 14 OECD countries. *Health Policy* 2016;120:1125-40.
  47. Mendelson A, Kondo K, Damberg C, Low A, Motúapuaka M, Freeman M, *et al.* The effects of pay-for-performance programs on health, health care use, and processes of care: A systematic review. *Ann Intern Med* 2017;166:341-53.
  48. Hong Y-R, Nguyen O, Yadav S, Etzold E, Song J, Duncan RP, *et al.* Early performance of hospital value-based purchasing program in medicare: A systematic review. *Med Care* 2020;58:734-43.
  49. Flodgren G, Eccles MP, Shepperd S, Scott A, Parmelli E, Beyer FR. An overview of reviews evaluating the effectiveness of financial incentives in changing healthcare professional behaviours and patient outcomes. *Cochrane Database Syst Rev* 2011;;2011:CD009255. doi: 10.1002/14651858.CD009255.
  50. Doran T, Kontopantelis E, Valderas JM, Campbell S, Roland M, Salisbury C, *et al.* Effect of financial incentives on incentivised and non-incentivised clinical activities: Longitudinal analysis of data from the UK quality and outcomes framework. *BMJ* 2011;342:d3590. doi: 10.1136/bmj.d3590.
  51. Mathes T, Pieper D, Morche J, Polus S, Jaschinski T, Eikermann M. Pay for performance for hospitals. *Cochrane Database Syst Rev* 2019;7:CD011156. doi: 10.1002/14651858.CD011156.pub2.