**International Journal of Accounting and Finance in Asia Pacific (IJAFAP)**
**Vol. 9 No. 1, pp. 288-304, Februari, 2026**
**E-ISSN: 2655-6502 P-ISSN: 2684-9763**
**https://www.ejournal.aibpmjournals.com/index.php/IJAFAP**

# Financial Incentive Reforms and Hospital Performance: A Systematic Review of Governance Mechanisms and Institutional Effects

**Merlin Hatumena[1*], Nyoman Sri Subawa[1], Ida Ayu Oka Martini[1], I Made Suidarma[1]**
[1]National University of Education, Denpasar, Indonesia
*Corresponding Email: hatumenamerlin@gmail.com

## ARTICLE INFORMATION

**Publication information**

**Research article**

**HOW TO CITE**
Hatumena, M., Subawa, N. S., Martini, I. A. O., & Suidarma, M. (2026). Financial incentive reforms and hospital performance: A systematic review of governance mechanisms and institutional effects. *International Journal of Accounting and Finance in Asia Pacific*, *9*(1), 288–304.

**DOI:**
**https://doi.org/10.32535/ijafap.v9i1.4442**



Received: 13 December 2025
Accepted: 19 January 2026
Published: 20 February 2026

## ABSTRACT

Hospital payment reforms, including pay-for-performance (P4P), diagnosis-related groups (DRG), bundled payments, value-based purchasing, and hybrid incentive systems, have been widely implemented to improve hospital quality, efficiency, and accountability, although empirical findings remain inconsistent across healthcare settings. This study aims to synthesize and evaluate the effectiveness of financial incentive reforms on hospital-level performance outcomes and identify the contextual factors influencing their implementation. Using a systematic review approach based on the PRISMA 2020 guidelines, this study analyzed 28 hospital-level studies published in Scopus-indexed journals (Q1–Q3) between 2015 and 2026. The findings indicate that 67.9% of the reviewed studies reported positive associations between payment reforms and hospital performance, particularly in process-based quality indicators, readmission outcomes, and selected efficiency measures. However, the effects on mortality, equity, utilization, and cost efficiency remain heterogeneous and context-dependent. The effectiveness of reforms is influenced by incentive magnitude, payment model type, indicator specificity, regulatory environment, baseline institutional performance, and digital monitoring capacity. These findings suggest that financial incentives are more effective when integrated into coherent governance and accountability systems.

**Keywords:** DRG; Healthcare Governance; Hospital Payment Reform; Hospital-Level Performance.

## INTRODUCTION

Over the past two decades, hospital payment reforms have become central instruments in global health system transformation. Governments and public payers have increasingly replaced volume-based reimbursement with performance-linked mechanisms such as pay-for-performance (P4P), diagnosis-related groups (DRG), bundled payments, and value-based purchasing. These reforms aim to improve quality, efficiency, and accountability in inpatient care while containing rising healthcare expenditures. The shift toward value-based competition reflects a broader policy agenda emphasizing measurable outcomes and financial alignment between providers and system-level objectives (Porter & Teisberg, 2006). Similarly, DRG-based systems have been promoted to enhance transparency and hospital efficiency across health systems (Busse et al., 2011). Consequently, hospital reimbursement structures are increasingly understood not merely as funding mechanisms but as governance instruments that shape organizational behavior.

Despite widespread implementation, empirical findings remain heterogeneous across outcome domains and regulatory contexts. Several studies report improvements in process-based quality indicators and reductions in readmissions (Girault et al., 2017; Gupta, 2021), yet evidence on mortality, long-term outcomes, and cost efficiency remains inconsistent and context-sensitive. Across the 28 empirical hospital-level studies synthesized in this review, 67.9% report statistically significant positive associations between financial incentives and performance outcomes, but effect magnitude and sustainability vary substantially across institutional environments. This divergence suggests that financial incentives do not operate uniformly as direct performance drivers; rather, their impact depends on structural and contextual conditions.

Three interrelated limitations in the existing literature help explain this inconsistency. First, most studies examine discrete programs within single-country contexts, limiting cumulative theoretical integration. Second, incentive reforms are predominantly framed through Agency Theory, emphasizing contractual alignment and information asymmetry between payers and providers (Eisenhardt, 1989; Jensen & Meckling, 1976), while underexploring the institutional environments that shape regulatory stability and organizational conformity (DiMaggio & Powell, 1983; Scott, 2014). Third, limited attention is given to heterogeneity in hospital-level capabilities particularly managerial competence, digital infrastructure, and internal coordination that may condition the translation of financial signals into sustained performance improvement (Barney, 1991; Wernerfelt, 1984). As a result, the field lacks an integrative framework capable of explaining cross-study heterogeneity in reform outcomes.

To address these gaps, this study conducts a systematic review of 28 empirical hospital-level studies published between 2015 and 2026. The study makes three contributions. First, it synthesizes fragmented empirical evidence across payment models and outcome domains within a unified analytical structure. Second, it reconceptualizes hospital payment reforms as multi-level governance instruments rather than isolated motivational mechanisms. Third, by integrating Agency Theory, Institutional Theory, and the Resource-Based View, it develops a structured explanatory framework clarifying how incentive design, institutional alignment, and organizational capacity interact to shape hospital-level performance outcomes. This governance-based perspective shifts the analytical focus from whether financial incentives “work” to the institutional and organizational conditions under which performance-linked reimbursement generates sustainable improvement.

## LITERATURE REVIEW

Hospital payment reforms are designed to realign provider incentives with system-level objectives of quality improvement, efficiency, and accountability. Major reform models include pay-for-performance (P4P), diagnosis-related groups (DRG), bundled payments, and value-based purchasing. Although these mechanisms differ in structure, they share a common premise: financial incentives can influence organizational behavior by linking reimbursement to measurable performance outcomes. Empirical research indicates that incentive-based reforms are most consistently associated with improvements in process-oriented indicators, including treatment adherence, documentation compliance, and selected readmission metrics. However, effects on mortality, long-term outcomes, and cost efficiency remain heterogeneous across institutional contexts.

The dominant explanatory framework in this literature is Agency Theory. From a principal–agent perspective, payment reform functions as a contractual mechanism intended to reduce information asymmetry and align hospital behavior with payer objectives (Jensen & Meckling, 1976; Eisenhardt, 1989). Incentive magnitude, monitoring intensity, and reward–penalty structure are expected to shape behavioral responsiveness. Under this logic, performance variation reflects differences in incentive salience and measurement credibility. While this framework explains how financial signals may alter organizational behavior, it does not fully account for persistent cross-country and cross-hospital heterogeneity observed in empirical findings.

Institutional Theory provides a complementary perspective by situating hospitals within broader regulatory and normative environments (DiMaggio & Powell, 1983; Scott, 2014). Hospitals operate under varying degrees of regulatory stability, enforcement capacity, professional norms, and legitimacy pressures. Financial incentives introduced into coherent and stable governance systems may reinforce accountability structures and generate substantive organizational change. Conversely, fragmented or unstable institutional environments may weaken incentive credibility, resulting in symbolic compliance or short-term adaptation. Institutional alignment therefore moderates the translation of financial signals into sustained performance improvement.

The Resource-Based View (RBV) further highlights heterogeneity in organizational capacity (Wernerfelt, 1984; Barney, 1991). Hospitals differ in managerial competence, digital infrastructure, data analytics capability, and internal coordination mechanisms. These resources shape the ability of organizations to respond strategically to performance-linked reimbursement. Incentive programs may amplify existing capacity differences, benefiting institutions equipped with robust information systems and integrated governance structures while generating limited gains among resource-constrained hospitals. From this perspective, financial incentives do not operate on uniform organizational foundations.

Taken together, these theoretical perspectives suggest that hospital payment reforms function not merely as motivational tools but as embedded governance mechanisms. Agency Theory explains incentive alignment at the contractual level; Institutional Theory accounts for regulatory credibility and normative pressures; and the Resource-Based View clarifies how internal capacity conditions implementation effectiveness. Yet existing empirical research rarely integrates these dimensions within a unified analytical structure. Most studies evaluate discrete programs within specific national contexts, limiting cumulative theoretical development and cross-study explanation.

This review addresses this gap by synthesizing contemporary hospital-level evidence within an integrated governance framework. By combining insights from Agency Theory,

Institutional Theory, and the Resource-Based View, the study advances a structured explanation for cross-study heterogeneity and clarifies the institutional and organizational conditions under which performance-linked reimbursement produces sustained improvements.

## RESEARCH METHOD

### Study Design

This study employs a systematic literature review to synthesize empirical evidence on hospital payment reforms and their effects on hospital-level performance. The review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) guidelines (Page et al., 2021). The objective was to identify, evaluate, and integrate contemporary empirical studies examining performance-linked reimbursement mechanisms in hospital settings.

### Search Strategy

A structured search was conducted in Scopus as the primary indexing database. The search covered publications from January 2015 to March 2026 to capture recent reform developments and contemporary evaluation designs. Search terms combined keywords related to hospital payment reform and performance outcomes, including “pay-for-performance,” “P4P,” “diagnosis-related groups,” “DRG,” “bundled payment,” “value-based purchasing,” “hospital performance,” “quality of care,” “readmission,” “mortality,” and “cost efficiency.”

The search was restricted to peer-reviewed journal articles published in English and indexed in Scopus (Q1–Q3). Reference lists of eligible articles were screened to identify additional relevant studies.

### Inclusion and Exclusion Criteria

Studies were included if they: (1) Employed quantitative, quasi-experimental, or observational empirical designs; (2) Examined hospital-level payment reforms (P4P, DRG, bundled payments, or value-based purchasing); (3) Reported measurable hospital performance outcomes (e.g., quality indicators, readmissions, mortality, cost efficiency, productivity, or equity-related outcomes); (4) Were published between 2015 and 2026 in peer-reviewed journals.

Studies were excluded if they: (1) Focused exclusively on individual physician compensation without hospital-level analysis; (2) Examined primary care or outpatient-only reforms; (3) Were purely conceptual, commentary-based, or lacked empirical analysis; (4) Addressed internal human resource incentives unrelated to external payment reform mechanisms.

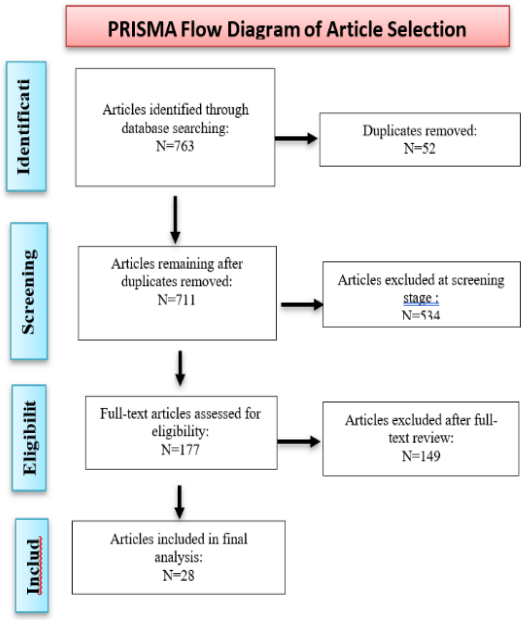
Following screening and eligibility assessment, 28 empirical hospital-level studies met the final inclusion criteria.

### Study Selection Process

The systematic identification of relevant research was executed in strict adherence to the PRISMA 2020 framework. Initially, a total of 763 records were aggregated through comprehensive electronic database queries. Following the removal of 52 redundant

entries, a refined pool of 711 citations was advanced to the preliminary screening phase. During the evaluation of titles and abstracts, 534 records were deemed ineligible and subsequently discarded. This led to a rigorous assessment of 177 full-text documents to determine their suitability. However, 149 articles were eliminated at this stage as they failed to align with the pre-defined inclusion parameters. Ultimately, 28 primary studies were consolidated into the final qualitative synthesis. The sequential progression encompassing identification, screening, eligibility appraisal, and final inclusion is visually detailed in the PRISMA flow chart (Figure 1). The diagram presents the stages of identification, screening, eligibility, and inclusion, resulting in 28 studies included in the final synthesis.

**Figure 1.** PRISMA 2020 flow diagram of the study selection process.




## Data Extraction and Synthesis

Data were extracted using a structured coding framework. Extracted variables included: (1) Country and institutional context; (2) Payment model type (P4P, DRG, bundled payment, value-based purchasing, hybrid mechanisms); (3) Study design and methodological approach; (4) Outcome domains (process quality, readmissions, mortality, cost efficiency, productivity, equity); (5) Direction and magnitude of reported effects; and (6) Identified contextual or moderating factors.

Given heterogeneity in institutional settings, outcome measures, and methodological designs, a quantitative meta-analysis was not conducted. Instead, the review employed a qualitative comparative synthesis approach, grouping findings by payment model and outcome domain.

## Analytical Approarch

The analysis proceeded in three stages. First, studies were categorized by payment model and primary outcome domain. Second, patterns of consistency and divergence were identified across institutional contexts. Third, findings were interpreted through the multi-level governance framework developed in the conceptual section, examining how

incentive design, institutional alignment, and organizational capacity contributed to observed heterogeneity.

## RESULTS

### Overview of Included Studies

The 28 included studies are summarized in Table 1. The sample encompasses diverse institutional contexts and payment reform models, including pay-for-performance (P4P), bundled payments, DRG-based systems, value-based purchasing, and hybrid incentive arrangements. Geographically, the evidence is concentrated in Taiwan (n = 10), the United States (n = 7), European health systems (n = 5), China (n = 4), Canada (n = 1), and South Korea (n = 1), reflecting evaluation activity in systems with established performance monitoring infrastructures.

**Table 1.** Summary of Included Studies (n = 28)

| Author (Year) | Country | Payment Model | Primary Outcome Domain | Effect Direction |
|---|---|---|---|---|
| (Baker et al., 2023) | USA | Bundled Payment (Medicare) | Cost / Efficiency | Positive |
| (Lu et al., 2021) | USA | P4P (Hospital QI) | Quality (Process) | Positive |
| (Cheng et al., 2023) | China | P4P | Digital Adoption / Quality | Positive |
| (Chopra et al., 2026) | USA | HRRP (Penalty-based P4P) | Readmission | Mixed |
| (Girault et al., 2017) | France | P4P | Quality (Process) | Mixed |
| (Gupta, 2021) | USA | HRRP | Readmission | Positive |
| (Han et al., 2024) | South Korea | Value-Based Payment | Infection Control | Positive |
| (Herbst et al., 2018) | Germany | P4P | Clinical Quality | Mixed |
| (Hsieh et al., 2016) | Taiwan | Diabetes P4P | Complications | Positive |
| (Hsieh et al., 2025) | Taiwan | Diabetes P4P | Health Outcomes / Utilization | Positive |
| (Huang et al., 2023) | Taiwan | HIV P4P | Treatment Initiation | Positive |
| (Kim et al., 2020) | USA | Penalty-based P4P | Surgical Quality | Mixed |
| (Liao et al., 2019) | USA | Bundled Payment | Patient Outcomes | Positive |
| (Lin C.Y. et al., 2022) | Taiwan | Team-Based Incentive | Quality | Positive |
| (Lin M. et al., 2023) | China | Performance-Based Incentive | Quality | Positive |
| (Lin T.Y. et al., 2016) | Taiwan | Diabetes P4P | Care Quality | Positive |
| (Lu et al., 2021) | Taiwan | P4P | Major Adverse Limb Events | Positive |
| (Núñez-Elvira et al., 2025) | UK | P4P | Access / Equity | Mixed |
| (Ramirez et al., 2016) | USA | Value-Based Purchasing | Quality | Positive |

| | | | | |
|---|---|---|---|---|
| (Slawomirski et al., 2024) | Multi-country | P4P | Patient Safety | Mixed |
| (Søgaard et al., 2015) | Denmark | Delegation-based Incentive | Efficiency | Mixed |
| (Tsai et al., 2023) | Taiwan | Diabetes P4P | Social Risk / Utilization | Mixed |
| (Vermeulen et al., 2016) | Canada | P4P | Waiting Time / Quality | Positive |
| (Yeh et al., 2025) | Taiwan | P4P | Hospitalization Risk | Positive |
| (Yen et al., 2022) | Taiwan | P4P | Long-term Outcomes | Positive |
| (Yen et al., 2025) | Taiwan | P4P | Mortality / CV Outcomes | Positive |
| (Zhao et al., 2022) | China | DRG | Performance Management | Mixed |
| (Zhou et al., 2025) | China | Bundled Payment | Referral Performance | Positive |

Across the sample, P4P constitutes the dominant reform model, particularly disease-specific programs in Taiwan (Hsieh et al., 2016; Lin T.Y. et al., 2016; Lu et al., 2021; Yeh et al., 2025; Yen et al., 2022; Yen et al., 2025; Hsieh et al., 2025; Tsai et al., 2023; Huang et al., 2023; Lin C.Y. et al., 2022), alongside broader quality-oriented initiatives in the United States and Europe (Chen et al., 2021; Gupta, 2021; Chopra et al., 2026; Kim et al., 2020; Girault et al., 2017; Herbst et al., 2018; Slawomirski et al., 2024; Núñez-Elvira et al., 2025; Vermeulen et al., 2016). Bundled payment models are examined primarily in the United States and China (Baker et al., 2023; Liao et al., 2019; Zhou et al., 2025), while DRG-based and governance-embedded systems are represented by Zhao et al. (2022) and Søgaard et al. (2015). Value-based purchasing mechanisms are evaluated in Ramirez et al. (2016) and Han et al. (2024).

**Publication Trends**

Publication activity increased after 2020, with the period 2021–2023 accounting for the largest proportion of included studies. Continued output during 2024–2026 indicates sustained scholarly attention to hospital payment reform evaluation.

**Table 2.** Distribution by Publication Period (n = 28)

| Publication Period | Number of Studies | Percentage |
|---|---|---|
| 2015–2017 | 6 | 21.4% |
| 2018–2020 | 3 | 10.7% |
| 2021–2023 | 11 | 39.3% |
| 2024–2026 | 8 | 28.6% |
| Total | 28 | 100% |

**Geographical Distribution**

The geographic distribution of studies reflects concentration in high-income and institutionally mature health systems. Taiwan represents the largest share, followed by the United States and European countries, with additional contributions from China, Canada, and South Korea.

**Table 3.** Distribution by Geographic Context (n = 28)

| Region / Country Group | Number of Studies | Percentage |
|---|---|---|
| United States | 7 | 25.0% |
| Taiwan | 10 | 35.7% |

| | | |
|---|---|---|
| China | 4 | 14.3% |
| Europe (UK, Germany, France, Denmark) | 5 | 17.9% |
| Canada | 1 | 3.6% |
| South Korea | 1 | 3.6% |
| Total | 28 | 100% |

**Distribution by Payment Model**

P4P constitutes 50.0% of the included studies (n = 14), followed by bundled payment models and value-based purchasing programs. DRG-based reimbursement and governance-embedded incentive systems represent a smaller but analytically relevant share.

**Table 4.** Distribution by Payment Model (n = 28)

| Payment Model | Number of Studies | Percentage |
|---|---|---|
| Pay-for-Performance (P4P) | 14 | 50.0% |
| Bundled Payment | 3 | 10.7% |
| DRG-Based Payment | 1 | 3.6% |
| Value-Based Purchasing (VBP) | 3 | 10.7% |
| Other Performance-Based / Hybrid Incentives | 7 | 25.0% |
| Total | 28 | 100% |

**Distribution by Primary Outcome Category**

Quality-related process indicators represent the most frequently examined outcome domain (n = 11), including treatment adherence, infection control, and care coordination (Han et al., 2024; Herbst et al., 2018; Lin C.Y. et al., 2022; Lin M. et al., 2023; Lin T.Y. et al., 2016; Huang et al., 2023; Slawomirski et al., 2024; Girault et al., 2017; Vermeulen et al., 2016; Ramirez et al., 2016).

Mortality and major clinical outcomes are examined primarily in diabetes-focused programs (Yen et al., 2025; Yen et al., 2022; Yeh et al., 2025; Hsieh et al., 2016; Lu et al., 2021; Hsieh et al., 2025), while readmission penalties are assessed in U.S.-based reforms (Gupta, 2021; Chopra et al., 2026). Cost and efficiency outcomes are concentrated in bundled payment and DRG-related evaluations (Baker et al., 2023; Liao et al., 2019; Zhao et al., 2022; Søgaard et al., 2015; Zhou et al., 2025). Equity and access dimensions are explored in Núñez-Elvira et al. (2025) and Tsai et al. (2023).

**Table 5.** Distribution by Primary Outcome Category (n = 36)

| Outcome Category | Number of Studies | Percentage |
|---|---|---|
| Quality of Care | 11 | 39.3% |
| Readmission / Utilization | 4 | 14.3% |
| Mortality / Major Clinical Outcomes | 6 | 21.4% |
| Cost & Efficiency | 5 | 17.9% |
| Productivity / Operational Performance | 2 | 7.1% |
| Total | 28 | 100% |

**Direction of Empirical Effects**

Across the 28 studies, 67.9% (n = 19) report statistically significant positive associations between financial incentive reforms and hospital-level performance outcomes. Mixed or context-dependent effects account for 17.9% (n = 5), while 14.3% (n = 4) report limited or statistically non-significant effects. Positive findings are most consistent in process-

based quality domains, whereas mortality, equity, and system-level cost outcomes display greater variability.

**Table 6.** Direction of Empirical Effects (n = 28)

| Effect Direction | Number of Studies | Percentage |
|---|---|---|
| Positive / Significant Improvement | 19 | 67.9% |
| Mixed / Conditional Effects | 5 | 17.9% |
| Limited / No Significant Effect | 4 | 14.3% |
| Total | 28 | 100% |

**Heterogenety and Moderating Factor**

Four recurring sources of heterogeneity emerge. First, incentive structure and financial salience influence responsiveness, particularly in bundled payment and DRG-based systems (Baker et al., 2023; Liao et al., 2019; Zhao et al., 2022; Zhou et al., 2025). Second, outcome measurability shapes effectiveness, with disease-specific programs demonstrating greater stability (Hsieh et al., 2016; Lin T. Y. et al., 2016; Lu et al., 2021; Yeh et al., 2025; Yen et al., 2022; Yen et al., 2025). Third, baseline institutional performance moderates the impact of reform implementation (Gupta, 2021; Chopra et al., 2026; Vermeulen et al., 2016). Fourth, governance embedding strengthens the sustainability of reforms, particularly in centralized healthcare systems (Hsieh et al., 2016; Lin T. Y. et al., 2016; Huang et al., 2023; Han et al., 2024).

**Table 7.** Moderating Factors Influencing the Effectiveness of Financial Incentives (n = 28)

| Moderating Dimension | Description | Observed Influence on Outcomes |
|---|---|---|
| Incentive Magnitude | Relative financial weight of reward/penalty | Stronger incentives associated with clearer performance responses |
| Payment Model Type | DRG, bundled payment, P4P, value-based payment | Bundled/DRG → stronger cost & efficiency effects; P4P → stronger process quality effects |
| Indicator Specificity | Narrow clinical targets vs. composite measures | Narrow targets yield more consistent improvements |
| Baseline Performance Level | High vs. low pre-reform performance | Greater gains observed among initially lower-performing hospitals |
| Regulatory Environment | Strength of monitoring and reporting systems | More structured systems show more stable improvements |
| Data Transparency | Public reporting and accountability mechanisms | Transparency enhances incentive responsiveness |

**Digital Infrastructure and Monitoring Capacity**

Digital capacity emerges as an enabling structural factor in several studies (Cheng et al., 2023; Han et al., 2024; Zhao et al., 2022). National claims integration and performance monitoring databases support evaluation in multiple Taiwanese programs (Hsieh et al., 2016; Lin T.Y. et al., 2016; Yen et al., 2022; Yen et al., 2025). Although not uniformly modeled, digital infrastructure appears to enhance measurement credibility and transparency.

**Table 8.** Digital Infrastructure and Monitoring Capacity (n = 28)

| Dimension | Number of Studies (n=36) | Percentage | Observed Role |
|---|---|---|---|

| | | | |
|---|---|---|---|
| Explicit digital/data infrastructure analysis | 8 | 28.6% | Improved monitoring accuracy and incentive credibility |
| Implicit data system reference | 6 | 21.4% | Supported accountability but limited structural integration |
| No digital consideration | 14 | 50.0% | Incentive effects evaluated without structural data-system analysis |
| Total | 28 | 100% | |

**Synthesis Interpretation**

Overall, the evidence indicates that financial incentives operate within layered governance configurations rather than as isolated economic instruments. Performance improvements are more stable when incentives are financially salient, targets are clearly defined, monitoring systems are credible, and reforms are embedded within coherent institutional frameworks. Cross-study heterogeneity therefore reflects variation in governance alignment rather than simple success or failure of incentive mechanisms.

**Risk of Bias Across Studies**

All 28 included studies met acceptable methodological standards under the structured appraisal described in the Methods section. Nevertheless, cross-study bias was assessed to evaluate the credibility and generalizability of the synthesized findings. Most studies employed quasi-experimental designs, including difference-in-differences, interrupted time series, and panel regression models. While these approaches provide stronger causal inference than cross-sectional analyses, they remain susceptible to unobserved confounding, spillover effects, and potential violations of parallel-trend assumptions. Randomized controlled trials were rare, reflecting practical constraints in evaluating large-scale payment reforms. As such, causal interpretations should be considered policy-informed rather than experimentally definitive.

Outcome measurement heterogeneity represents a second source of potential bias. Process-based indicators generally demonstrate higher measurement stability, whereas mortality, cost efficiency, and utilization outcomes are influenced by case-mix variation and broader system-level factors. Although many studies adjusted for hospital characteristics and patient risk profiles, adjustment strategies were not fully uniform, limiting strict cross-study comparability.

Contextual bias also warrants consideration. The majority of included studies were conducted in high-income health systems with advanced reimbursement infrastructure and integrated data environments. This concentration may limit generalizability to lower-resource settings. In addition, evaluations of nationwide reforms may confound incentive effects with concurrent structural changes in health system governance.

Publication and reporting bias cannot be excluded, as the review included only peer-reviewed Scopus-indexed (Q1–Q3) articles. Positive or policy-relevant findings may therefore be overrepresented. However, the presence of mixed (17.9%) and non-significant (14.3%) results suggests that favorable effects are not uniformly overstated.
Selective emphasis on incentivized indicators was observed in several studies, with limited systematic assessment of unintended consequences such as service shifting, risk selection, or administrative burden. These outcomes were occasionally discussed but not consistently measured.

**Table 9.** Cross-Study Risk of Bias Assessment (n = 28)

| Bias Dimension | Source of Potential Bias | Observed Pattern | Overall Risk Level |
|---|---|---|---|
| Study Design Bias | Quasi-experimental predominance | Strong statistical control but limited randomization | Moderate |
| Measurement Bias | Variability in outcome definitions | Process indicators more stable than mortality/cost outcomes | Moderate |
| Contextual Bias | High-income country concentration | Limited LMIC representation | Moderate |
| Publication Bias | Peer-reviewed Scopus-indexed journals only | Possible underreporting of null effects | Low–Moderate |
| Policy Confounding | Simultaneous systemic reforms | Difficulty isolating pure incentive effects | Moderate |

Overall, the evidence base demonstrates acceptable methodological rigor with a moderate cross-study risk profile. No systematic high-risk bias was identified that would invalidate the primary synthesized conclusions. However, the predominance of quasi-experimental designs, outcome heterogeneity, and concentration in advanced health systems indicate that findings should be interpreted as context-sensitive rather than universally generalizable.

## DISCUSSION

### Principal Finding

This review synthesizes 28 empirical studies examining hospital-level financial incentive reforms and their impact on institutional performance. Across the literature, improvements are most consistently observed in process-based quality indicators and selected readmission outcomes. In contrast, evidence concerning mortality reduction and cost-efficiency remains heterogeneous and strongly shaped by contextual conditions.

These findings suggest that financial incentives tend to operate most effectively as mechanisms that align organizational behavior with specific performance targets rather than as instruments capable of producing comprehensive institutional transformation. Programs linked to clearly defined and operationally measurable indicators, supported by credible monitoring and reporting systems, appear more likely to generate observable performance responses. In this context, financial incentives primarily influence domains in which performance expectations are explicitly articulated and routinely evaluated.

### Interpretation of Effect Patterns

The predominance of positive findings in process-based indicators reflects the structural clarity embedded in many incentive programs. Clearly specified and operationally measurable targets reduce interpretive ambiguity and strengthen organizational responsiveness to financial signals. When performance indicators are directly observable and systematically monitored, hospitals are better able to translate incentive structures into managerial adjustments and clinical practice improvements.

By contrast, outcomes such as mortality rates and cost-efficiency are shaped by a broader constellation of structural determinants, including case-mix complexity,

technological capacity, workforce composition, and system-level resource allocation. Because these factors extend beyond the immediate scope of incentive design, financial mechanisms alone may exert limited influence on such multidimensional outcomes. Variability in these domains therefore reflects the complexity of healthcare performance measurement rather than the inherent ineffectiveness of incentive-based reforms.

Several studies also report modest effect sizes and notable variation across institutional settings. While systematic evidence of performance deterioration remains limited, the literature highlights potential unintended responses to incentive structures. These include service shifting, selective patient avoidance, administrative burden, and short-term optimization strategies aimed primarily at meeting measurable targets. Such findings underscore the importance of balanced indicator selection and the need for governance mechanisms capable of mitigating strategic behavior associated with performance measurement systems.

**Incentive Design, Institutional Context, and Organizational Capacity**

The heterogeneity observed across studies further emphasizes the role of program architecture and institutional context in shaping policy outcomes. Incentive magnitude, indicator specificity, and the credibility of reporting systems consistently influence the extent to which hospitals respond to performance-based payment reforms. Hospitals operating within stable regulatory environments and supported by structured monitoring systems tend to demonstrate more consistent and durable improvements.

Evidence also suggests the presence of diminishing marginal returns among already high-performing institutions. In such contexts, additional financial incentives may produce relatively limited incremental improvements because performance levels are already approaching established benchmarks. Conversely, hospitals with lower baseline performance often demonstrate greater relative gains following the introduction of incentive programs. This pattern indicates that financial incentives frequently function as corrective alignment mechanisms that encourage improvement among underperforming institutions rather than as universal performance accelerators.

Organizational capacity represents another critical moderating factor. Hospitals equipped with established performance management systems, standardized clinical protocols, and coordinated administrative structures are better positioned to translate financial incentives into operational change. Where such capacities are limited, the influence of payment reform may be constrained by implementation challenges and fragmented internal coordination.

In addition, the geographic concentration of existing empirical evidence in high-income health systems highlights a structural pattern within the literature. Differences in administrative capability, regulatory stability, and data infrastructure may significantly influence how financial incentives operate across health system environments. As a result, policy responses to payment reform may vary substantially across institutional and national contexts.

**Theoretical Contribution: A Multi-Level Governance Perspective**

This study contributes to the literature by conceptualizing hospital payment reform as part of a broader multi-level governance architecture rather than solely as a motivational policy instrument. Drawing on Agency Theory (Jensen & Meckling, 1976; Eisenhardt, 1989), Institutional Theory (DiMaggio & Powell, 1983; Scott, 2014), and the Resource-Based View (Wernerfelt, 1984; Barney, 1991), the findings indicate that the effectiveness of financial incentives is conditioned by governance alignment across multiple institutional layers.

At the micro level, Agency Theory suggests that performance-linked reimbursement mechanisms seek to align the behavior of healthcare professionals with organizational objectives. However, the effectiveness of such alignment depends heavily on clearly defined performance indicators and credible monitoring systems. Without reliable evaluation frameworks, financial incentives may encourage symbolic compliance rather than substantive behavioral change.

At the organizational level, internal governance structures shape how incentive programs are interpreted and implemented. Hospitals characterized by coordinated management systems, embedded performance monitoring routines, and standardized clinical pathways are generally better positioned to respond effectively to payment reforms. In these settings, financial incentives operate not as isolated policy tools but as components integrated within broader organizational governance arrangements.

At the macro level, Institutional Theory helps explain cross-country variation in policy outcomes. Health systems characterized by regulatory stability, coherent reporting frameworks, and mature institutional structures tend to demonstrate stronger responsiveness to performance-based incentives. Conversely, fragmented governance environments and policy volatility may weaken the credibility and sustainability of reform initiatives.

Across these governance levels, digital infrastructure emerges as an enabling capability consistent with the Resource-Based View (Wernerfelt, 1984; Barney, 1991). Integrated health information systems improve measurement precision, reduce information asymmetry, and strengthen accountability mechanisms across organizations and regulatory authorities. Consequently, the effectiveness of financial incentive programs depends not only on payment design but also on the institutional and organizational capacities that support performance measurement and governance coordination.

**Policy Implications**

The findings indicate that payment reform effectiveness depends less on the presence of incentives than on their integration within coherent governance systems. Financial incentives generate more consistent responses when supported by measurable indicators, financially salient reward structures, transparent monitoring, and organizational coordination mechanisms. Strengthening data infrastructure and embedding performance management within institutional routines are therefore central to maximizing reform impact. Payment adjustment without governance integration is likely to produce fragmented and uneven outcomes.

## CONCLUSION

This systematic review synthesizes evidence from 28 empirical studies evaluating hospital-level financial incentive reforms, including pay-for-performance, bundled payments, diagnosis-related group systems, and value-based reimbursement models. Across these studies, financial incentives consistently improve process-based quality indicators and selected readmission outcomes, while effects on mortality, cost-efficiency, and equity remain heterogeneous and context-dependent. The effectiveness of financial incentives cannot be attributed to payment design alone. Performance outcomes are strongly shaped by the alignment of incentive magnitude, indicator specificity, monitoring credibility, organizational capacity, and institutional context. Hospitals operating within stable regulatory environments and supported by robust data infrastructure exhibit more consistent and durable responses to performance-linked reimbursement. Lower-

performing hospitals generally experience greater relative gains, highlighting the corrective potential of incentive programs when embedded appropriately.

Conceptually, integrating Agency Theory, Institutional Theory, and the Resource-Based View illustrates that hospital financial incentives function as multi-level governance instruments rather than isolated motivators. Their effectiveness depends on coherence across micro-level accountability mechanisms, organizational integration, and macro-level institutional stability, with digital infrastructure acting as a structural enabler for measurement credibility and performance transparency. Overall, financial incentives alone are insufficient as stand-alone policy tools. Sustainable improvements in hospital performance require embedding payment reforms within coherent governance architectures that support institutional readiness, credible monitoring, and coordinated organizational processes, ensuring that incentive structures translate into measurable and durable health system outcomes.

## LIMITATION

Several limitations should be considered when interpreting the findings of this review. First, the predominance of quasi-experimental designs in the included studies limits definitive causal inference, as randomization was rarely employed and residual confounding cannot be fully excluded. Second, the majority of evidence originates from high-income and institutionally mature health systems, particularly Taiwan, the United States, and Europe, which constrains the generalizability of results to low- and middle-income countries or resource-constrained settings. Third, heterogeneity in outcome measurement, particularly for mortality, cost-efficiency, and equity, restricts strict cross-study comparability and may contribute to variability in observed effect sizes. Fourth, selective reporting and publication bias cannot be entirely ruled out, as only peer-reviewed, Scopus-indexed studies were included. Finally, although this review integrates Agency Theory, Institutional Theory, and the Resource-Based View, the reliance on secondary literature limits the ability to directly observe organizational mechanisms and contextual adaptations in practice.

## ACKNOWLEDGMENT

The authors wish to express their sincere gratitude to the institutions and researchers whose studies were included in this review. We also thank colleagues and reviewers who provided valuable feedback during the preparation of this manuscript, contributing to the clarity and rigor of the analysis.

## DECLARATION OF CONFLICTING INTERESTS

The authors have declared no potential conflicts of interest concerning the study, authorship, and/or publication of this article.

## REFERENCES

Baker, M. C., Hahn, E. N., Dreyer, T. R., & Horvath, K. A. (2023). Succeeding in Medicare's newest bundled payment program: Results from teaching hospitals. *Healthcare, 11*(1), 100672. https://doi.org/10.1016/j.hjdsi.2022.100672

Barney, J. B. (1991). Firm resources and sustained competitive advantage. *Journal of Management, 17*(1), 99–120. https://doi.org/10.1177/014920639101700108

Busse, R., Geissler, A., Quentin, W., & Wiley, M. (2011). Diagnosis-related groups in Europe: Moving towards transparency, efficiency and quality in hospitals. *BMJ, 343*, d4537. https://doi.org/10.1136/bmj.d4537

Cheng, N., Li, H., & Bang, Y. (2023). Pay-for-performance schemes and hospital HIT adoption. *Decision Support Systems, 164*, 113868. https://doi.org/10.1016/j.dss.2022.113868

Chopra, Z., Ryan, A. M., & Hoffman, G. J. (2026). Hospital Readmission Reduction Program penalties for hospitals with high Medicare Advantage penetration. *JAMA Network Open, 9*(1), e2554972.

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review, 48*(2), 147–160. https://doi.org/10.2307/2095101

Eisenhardt, K. M. (1989). Agency theory: An assessment and review. *Academy of Management Review, 14*(1), 57–74. https://doi.org/10.5465/amr.1989.4279003

Girault, A., Bellanger, M., Lalloué, B., Loirat, P., Moisdon, J. C., & Minvielle, E. (2017). Implementing hospital pay-for-performance: Lessons learned from the French pilot program. *Health Policy, 121*(4), 407–417. https://doi.org/10.1016/j.healthpol.2017.02.002

Gupta, A. (2021). Impacts of performance pay for hospitals: The Readmissions Reduction Program. *American Economic Review, 111*(4), 1241–1283. https://doi.org/10.1257/aer.20181504

Han, K. T., Kim, S., Kim, G. O., Lee, S., & Kwon, Y. U. (2024). Quality control efforts of medical institutions: The impacts of a value-based payment system on medical staff and healthcare-associated infections. *Journal of Hospital Infection, 153*, 3–13.

Herbst, T., Foerster, J., & Emmert, M. (2018). The impact of pay-for-performance on the quality of care in ophthalmology: Empirical evidence from Germany. *Health Policy, 122*(6), 667–673. https://doi.org/10.1016/j.healthpol.2018.04.014

Hsieh, H. M., Lin, T. H., Lee, I. C., Huang, C. J., Shin, S. J., & Chiu, H. C. (2016). The association between participation in a pay-for-performance program and macrovascular complications in patients with type 2 diabetes in Taiwan. *Preventive Medicine, 85*, 53–59. https://doi.org/10.1016/j.ypmed.2016.02.035

Hsieh, H. M., Wang, Y. H., & Chen, H. F. (2025). Associations between participation in a diabetes pay-for-performance program and health outcomes and healthcare utilization among people with comorbid schizophrenia and type 2 diabetes in Taiwan. *General Hospital Psychiatry, 94*, 99–107.

Huang, Y. F., Pan, L. C., Yang, J. Y., Liao, Y. H., Su, H. J., Mei, N. H., & Tsai, Y. C. (2023). Assessment of performance regarding confirmatory diagnosis and initiation of antiretroviral therapy under a modified national HIV testing algorithm and pay-for-performance program in Taiwan. *Journal of Microbiology, Immunology and Infection, 56*(6), 1139–1146. https://doi.org/10.1016/j.jmii.2022.10.007

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics, 3*(4), 305–360. https://doi.org/10.1016/0304-405X(76)90026-X

Kim, K. M., Max, W., White, J. S., Chapman, S. A., & Muench, U. (2020). Do penalty-based pay-for-performance programs improve surgical care more effectively than other payment strategies? A systematic review. *Annals of Medicine and Surgery, 60*, 623–630. https://doi.org/10.1016/j.amsu.2020.02.010

Liao, J. M., Emanuel, E. J., Venkataramani, A. S., Huang, Q., Dinh, C. T., Shan, E. Z., & Navathe, A. S. (2019). Association of bundled payments for joint replacement surgery and patient outcomes with simultaneous hospital participation in accountable care organizations. *JAMA Network Open, 2*(9), e1912270. https://doi.org/10.1001/jamanetworkopen.2019.12270

Lin, C. Y., Tsai, Y. H., & Chen, J. L. (2022). Team-based incentive programs and healthcare quality: A case study in Taiwan. *Health Policy and Technology, 11*(1), 100584. https://doi.org/10.1016/j.hlpt.2021.100584

Lin, M., Chen, Z., & Huang, W. (2023). Performance-based incentives and quality outcomes in public hospitals: Evidence from panel data analysis. *Health Economics Review, 13*(1), 45. https://doi.org/10.1186/s13561-023-00441-8

Lin, T. Y., Chen, C. Y., Huang, Y. T., Ting, M. K., Huang, J. C., & Hsu, K. H. (2016). The effectiveness of a pay-for-performance program on diabetes care in Taiwan: A nationwide population-based longitudinal study. *Health Policy, 120*(11), 1313–1321. https://doi.org/10.1016/j.healthpol.2016.09.011

Lu, C. W., Wu, Y. F., Chen, T. H., Chung, C. M., Lin, C. L., Lin, Y. S., & Lin, M. S. (2021). A nationwide cohort investigation on pay-for-performance and major adverse limb events in patients with diabetes. *Preventive Medicine, 153*, 106787. https://doi.org/10.1016/j.ypmed.2021.106787

Núñez-Elvira, A., Feng, Y., Kristensen, S. R., Lorgelly, P., Meacock, R., Siciliani, L., & Sutton, M. (2025). Does pay for performance affect socioeconomic inequalities in access? Evidence from hospital specialised care in England. *Health Policy*.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., & Mulrow, C. D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71

Porter, M. E., & Teisberg, E. O. (2006). *Redefining health care: Creating value-based competition on results*. Harvard Business School Press.

Ramirez, A. G., Tracci, M. C., Stukenborg, G. J., Turrentine, F. E., Kozower, B. D., & Jones, R. S. (2016). Physician-owned surgical hospitals outperform other hospitals in Medicare value-based purchasing program. *Journal of the American College of Surgeons, 223*(4), 559–567. https://doi.org/10.1016/j.jamcollsurg.2016.07.733

Scott, W. R. (2014). *Institutions and organizations: Ideas, interests, and identities* (4th ed.). Sage Publications.

Slawomirski, L., Hensher, M., Campbell, J., & de Graaff, B. (2024). Pay-for-performance and patient safety in acute care: A systematic review. *Health Policy, 143*, 105051. https://doi.org/10.1016/j.healthpol.2024.105051

Søgaard, R., Kristensen, S. R., & Bech, M. (2015). Incentivising effort in governance of public hospitals: Development of a delegation-based alternative to activity-based remuneration. *Health Policy, 119*(8), 1076–1085. https://doi.org/10.1016/j.healthpol.2015.06.005

Tsai, W. C., Huang, K. H., Chen, P. C., Chang, Y. C., Chen, M. S., & Lee, C. B. (2023). Effects of individual and neighborhood social risks on diabetes pay-for-performance program under a single-payer health system. *Social Science & Medicine, 326*, 115930. https://doi.org/10.1016/j.socscimed.2023.115930

Vermeulen, M. J., Stukel, T. A., Boozary, A. S., Guttmann, A., & Schull, M. J. (2016). The effect of pay for performance in the emergency department on patient waiting times and quality of care in Ontario, Canada. *Annals of Emergency Medicine, 67*(4), 496–505. https://doi.org/10.1016/j.annemergmed.2015.10.012

Wernerfelt, B. (1984). A resource-based view of the firm. *Strategic Management Journal, 5*(2), 171–180. https://doi.org/10.1002/smj.4250050207

Yeh, S. C., Chen, Y. H., Van Velzen, R., & Lin, P. H. (2025). The climate change literacy of public officials in Taiwan: Implications and strategies for global adaptation. *Policy Studies, 46*(2), 168–201. https://doi.org/10.1080/01442872.2024.2304086

Yen, F. S., Wei, J. C. C., Hung, Y. T., Hsu, C. C., & Hwu, C. M. (2022). Long-term outcomes of the pay-for-performance program for patients with young-onset (20–40 years of age) type 2 diabetes. *Diabetes Research and Clinical Practice, 193*, 110136. https://doi.org/10.1016/j.diabres.2022.110136

Yen, F. S., Wei, J. C. C., Lin, S. Y., Cho, J. D. Y., Chi, M. H., & Hsu, C. H. (2025). The impact of pay-for-performance care on the mortality and cardiovascular outcomes in older adults with newly diagnosed type 2 diabetes: A nationwide population-based cohort study. *Journal of the American Medical Directors Association, 26*(2), 105382. https://doi.org/10.1016/j.jamda.2024.105382

Zhao, S., Gu, Y., & Huang, Z. (2022). Building a performance management system for hospitals based on DRG payment. *Journal of Sensors, 2022*, 7001423. https://doi.org/10.1155/2022/7001423

Zhou, W., Gan, Y., & Dang, Y. (2025). Bundled vs. unbundled: Impact of payment scheme on performance in a healthcare downstream referral program. *Journal of Management Science and Engineering, 10*(2), 143–157. https://doi.org/10.1016/j.jmse.2024.10.004

## ABOUT THE AUTHOR(S)

**1st Author**
Merlin Hatumena is an undergraduate student at Universitas Pendidikan Nasional, Denpasar, Indonesia. His academic interests include healthcare management, hospital performance, and health policy. Email: hatumenamerlin@gmail.com

**2nd Author**
Nyoman Sri Subawa is a lecturer at Universitas Pendidikan Nasional, Denpasar, Indonesia. His research focuses on management studies, organizational governance, and healthcare systems. Email: shribawa@undiknas.ac.id. Orchid ID: https://orcid.org/0000-0001-5837-4863

**3rd Author**
Ida Ayu Oka Martini is a lecturer at Universitas Pendidikan Nasional, Denpasar, Indonesia. Her academic work includes hospital management, performance evaluation, and public health administration. Email: dayuokamartini@undiknas.ac.id

**4th Author**
I Made Suidarma is a lecturer at Universitas Pendidikan Nasional, Denpasar, Indonesia. His research interests cover healthcare policy, organizational development, and hospital performance management. Email: suidarma@undiknas.ac.id Orchid ID: https://orcid.org/0009-0002-3198-1801