



HHS Public Access

Author manuscript

Health Aff (Millwood). Author manuscript; available in PMC 2024 June 01.

Published in final edited form as:

Health Aff (Millwood). 2023 June ; 42(6): 813–821. doi:10.1377/hlthaff.2022.01455.

Value-based purchasing design and effect: a systematic review and analysis

Abhinav Pandey, BA¹, Diana Eastman, MS², Heather E. Hsu, MD MPH³, Michaela Kerrissey, PhD⁴, Meredith B. Rosenthal, PhD⁴, Alyna T. Chien, MD, MS^{5,6}

¹Weill Cornell Medical College, New York, New York, USA.

²Open Health Solutions, The MITRE Corporation, Bedford, MA.

³Department of Pediatrics, Boston University School of Medicine, Boston, MA.

⁴Harvard T. Chan School of Public Health, Boston, MA.

⁵Department of Pediatrics, Harvard Medical School, Boston, MA.

⁶Division of General Pediatrics, Boston Children's Hospital, Boston, MA.

Abstract

Over the past two decades in the United States (US), all major payer types—commercial, Medicare, Medicaid, and multi-payer coalitions—have introduced value-based purchasing (VBP) contracts to reward providers for improving health care quality while reducing spending. This systematic review qualitatively characterizes the financial and non-financial features of VBP programs and examines how such features combine to create a level of program intensity that relates to desired quality and spending outcomes. Higher-intensity VBP programs are more frequently associated with desired quality processes, utilization measures, and spending reductions than lower-intensity programs. Thus, while there may be reasons for payers and providers to opt for lower-intensity programs (e.g., to increase voluntary participation), these choices apparently have consequences for spending and quality outcomes.

Keywords

healthcare; incentives; value-based purchasing; quality; spending; global contracts; bundled payments

INTRODUCTION

In the early 2000s in the United States (US), major payers introduced value-based purchasing (VBP) contracts or programs that entailed incentives for care quality (i.e., pay for performance) without direct relation to spending.¹ By the end of that decade, VBP evolved into contracts that blended incentives for improving quality with those for reducing spending.^{2–5}

CORRESPONDING AUTHOR & CONTACT INFORMATION: Abhinav Pandey (abp4003@med.cornell.edu), Weill Cornell Medical College, New York, New York.

Numerous systematic reviews highlight the variation of VBP programs in terms of the depth and breadth of their performance metrics and surmise that VBP has delivered modest reductions in spending but had null effects on quality.^{6–16} However, these prior systematic reviews have limited their focus to certain sponsors (e.g., Medicare only or commercial payers only) or types of VBP incentives (e.g., pay-for-performance only, bundled payment initiatives only, or a blend of pay-for-performance and accountable care organizations),^{9–16} and are agnostic to the potential role played by overarching VBP program design – that is, the choices that are made about the mix of financial and non-financial features that, together, result in differing levels of program intensity. However, because VBP programs are implemented across multiple sponsor types and with a variety of financial and non-financial features, it is likely that heterogeneity in VBP design has meaningful implications.³ Identifying and characterizing VBP design choices and incorporating such choices into our understanding of the burgeoning research on VBP may move the field toward an understanding of the impact of VBP program design on outcomes.

The financial aspects of VBP include bonuses, penalties, and financial risk-sharing arrangements that can be operationalized in a variety of ways (Appendix-Figure IIA).¹⁷ VBP contracts may introduce financial incentives to reduce spending in any of several ways: as a resource efficiency measure within a pay-for-performance framework (e.g., providing bonuses to providers who succeed in reducing unnecessary hospital admissions);¹⁸ as a prospective or bundled payment focused on a particular episode of service (e.g., a single, comprehensive payment for the series of services that constitute a total knee replacement); or, as responsibility for total medical expenditures tied to financial risk that can be upside-only (i.e., not imposing losses for exceeding budget but offering payment if savings are generated) or two-sided (i.e., making providers liable for exceeding budget and offering payment if savings are generated).¹⁹

VBP contracts may also tie quality-improvement incentives to spending reduction incentives in myriad ways, such as through pay-for-performance bonuses (as noted above), by setting minimum quality standards that must be met to share in savings^{20–22} (i.e., gates), or by using quality performance levels to adjust the size of shared savings or losses (i.e., ladders).^{23–25} See Appendix-Table IIA for a glossary of key terms related to financial and quality incentives.¹⁷

VBP contracts often have non-financial supports, which have received less attention in the literature (Appendix-Section IIB).^{3,17} Non-financial supports have the potential to make it easier for providers to respond to the spending and quality incentives that VBP programs introduce.²⁶ The non-financial supports that we have encountered in our past work include the provision of: (1) analyzed data, reports, or lists to help health care providers take action (e.g., list of patients needing immunizations so that they can be contacted²⁷ or the building of reporting platforms²⁸); (2) technical assistance in the form of leadership or change management training for personnel (e.g., learning collaboratives²⁹); (3) infrastructure payments or funds to add new personnel or retool existing staff or their skills (e.g., delivery system reform incentive payments [DSRIP] used in Medicaid^{30–32}); (4) raw claims data to allow providers to investigate care and spending patterns (e.g., medical expenditures by categories³³); (5) risk management support (e.g., stop-loss that caps organizations' exposure

to penalties should extremely high spending occur³⁰); and, (6) care management support, such as new personnel (e.g., patient navigators, community health workers³¹), who can organize care across all settings of the broader health care system. Non-financial supports may be critical for desired VBP effects,^{27,34} especially for providers that have traditionally operated within siloes with little visibility into care that occurs in upstream or downstream entities.

Although past research has tended to focus on individual VBP design features in isolation, in reality VBP programs use an array of financial incentives for spending reduction and quality improvement alongside non-financial supports—and they do so with substantive variation. Spending and quality incentives can vary in their strength. Non-financial supports can vary with respect to the number and types used. Delivered together, incentives and supports may both add to one another and interact, such that non-financial supports further spur the ability to reach spending and quality targets. To identify the potentially additive or interactive aspects of these features requires the ability to characterize VBP program intensity across three dimensions—spending and quality incentives and non-financial supports—and to systematically plot specific programs along them, relative to desired program outcomes.

To our knowledge, this study is the first to combine a systematic review of VBP programs with the capture and characterization of their financial and non-financial features, to combine VBP features to assess program intensity, and to relate program intensity to quality and spending effects.^{3,6–15} This perspective is vital to informing policymakers as VBP enters its third decade as the cornerstone of US payment reform. In reality, policymakers must make choices not only about individual features of VBP but also about the intensity and mix of features to promote overall. To provide empirical support for such choices, we:

1. Systematically review empirically evaluated VBP programs that have introduced both quality improvement and spending reduction incentives;
2. Characterize both the incentives and supports used in these VBP programs and combine these two dimensions to array programs into higher- and lower-intensity designs; and,
3. Examine the degree to which the intensity of VBP program design relates to desired spending reduction and quality improvement outcomes.

STUDY DATA AND METHODS

Study Overview

Because the range of VBP program designs and their corresponding outcomes were too heterogeneous to conform to a formal meta-analysis, we conducted this study by combining a traditional systematic review with a qualitative analysis in three parts (Appendix).¹⁷ The first part was a systematic review that followed Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. The second part was a program-level qualitative analysis of the financial and non-financial features of VBP programs, analyzing variation in each dimension to create a construct for VBP program intensity and then identifying programs that were delivering higher- versus lower-intensity incentives. The

third part was a program-level qualitative analysis of how higher- versus lower-intensity VBP programs related to spending and quality findings across each program's empirical evaluations. All parts involved independent review by multiple study team members and multiple rounds of discussion until consensus was reached.

Identification of Empirically Evaluated VBP Programs

We included VBP programs that introduced incentives for reducing spending and improving care quality, were based in the US, and had been empirically evaluated (i.e., had been studied to evaluate an outcome directly tied to the VBP program's intended effect) (Appendix-Section I).¹⁷ We searched medical library databases (Medline[®] and the Cochrane Library) for English-language studies published between January 1, 2000 and December 31, 2017, and rechecked through July 1, 2020 to capture evaluations of VBP programs that had been in place for up to four years. The search for VBP program descriptions occurred between August 1, 2017, and September 30, 2018. We assessed the rigor of identified studies using the Downs and Black tool,³⁵ which was modified to discriminate the rigor of quasi-experimental designs.

Characterization of VBP Program Design: Higher- Versus Lower-Intensity Features

In the absence of empirical information about how VBP programs were perceived by participants or acted upon, we followed a conceptual framework to express how the spending and quality incentives of VBP programs could differ in strength. For example, using a pay-for-performance measure to reward resource efficiency was considered to be weaker than two-sided financial risk (Appendix-Section II).¹⁷ However, it was more difficult to decide *a priori* which types of non-financial supports were stronger than others, and empirically some programs had a few in place while others had more, so we simply counted the total number of non-financial supports. We then combined the information from the three dimensions. VBP programs were designated as higher-intensity if: (a) spending incentives involved upside-only or two-sided risk; (b) spending incentives were tied to quality performance via gates or ladders; and (c) non-financial supports included at least four of six possible types (Appendix-Section IIA).^{17,36}

Characterization of Program-Level VBP Outcomes

While designated study team members were characterizing VBP program design, others were independently characterizing the spending and quality outcomes from each empirical evaluation (Appendix-Section IIIA, Appendix-Section 1D).¹⁷ For each empirical evaluation, spending or quality outcomes were characterized as positive if outcomes were statistically different in the desired direction (e.g., down for spending, up for quality) relative to comparisons at the 0.05 level, negative if they moved in the opposite direction, and null if there was no significant difference.

As most evaluations studied several measures, the overall outcome was defined as positive if a majority of evaluated measures trended favorably compared to controls at a significant level. Similarly, if a majority of evaluated measures trended negatively, the overall result was considered to be negative. However, as 15 VBP programs had been evaluated multiple times, and findings could differ across evaluations, if necessary, VBP outcomes were also

summarized to the program level and characterized as mixed-positive or mixed-negative when individual empirical evaluations differed (i.e., one empirical evaluation found a positive effect while another yielded a null one) or when there was a majority of null findings combined with a plurality of positive or negative findings (e.g., a study assessing 5 quality measures that found no change in 3 measures but found a positive directional change in 2 would be defined as mixed-positive) (Appendix-Section IIIA, Appendix-Section ID).¹⁷

Relationship Between VBP Program Intensity and Program-Level Outcomes

The qualitative analysis of the relationship between VBP program intensity and spending and quality outcomes was pursued by analyzing how the size of the literature and balance of outcomes shifted across strata related to literature rigor and program intensity criteria (Appendix-Section IIIB).¹⁷ For each spending or quality outcome type of interest, strata were: (a) All programs; (b) Rigor >14, which presents VBP program findings when studies were restricted to those with modified Downs and Black scores of 14 and above; (c) Higher-intensity programs; and (d) Lower-intensity programs. Strata (a), (c), and (d) included all studies irrespective of study rigor level.

LIMITATIONS

This systematic review relied on publicly available information on program design and effect. Documented information may reflect publication, social desirability, and recall biases. It usually lacks details regarding the subjective experiences of organizations (e.g., how much they feel they can influence processes and outcomes), their payer mix (e.g., the proportion of patients in VBP contracts), the alignment of the incentives within the VBP contract versus other contracts, or the degree to which VBP incentives delivered to organizations is transmitted to frontline physicians. However, the programmatic features identified in this analysis reflect the documented contract features thought, by payers and providers, to be important in VBP design.²⁶

RESULTS

This systematic review identified 24 unique VBP programs that incentivized spending reduction and quality improvement. These 24 programs had been evaluated in 69 unique publications with a median of two publications per program (mean 2.88, standard deviation 2.3, maximum 10) (Exhibit 1). In terms of the types of outcomes studied, all 24 programs had been evaluated for their effects on quality in 64 unique publications, whereas 22 of the 24 programs had been evaluated for effects on spending in 44 unique publications. The level of rigor of the studies evaluating quality improvement and spending appeared comparable (Appendix-Section 1C).¹⁷ The median modified Downs and Black scores for studies of quality improvement was 17.8 (range 5.5–25.5); for studies of spending reduction, it was 17 (range 5.5–26). Eight VBP programs were sponsored by Medicare, two by Medicaid, eight by all- or multi-payer groups, and six by commercial health plans (Appendix-Table IIB–c).¹⁷

All but one of the 24 VBP programs had published information about the non-financial supports accompanying the financial and quality incentives. The median number of non-financial supports was four, with two programs offering all six types of non-financial

supports. The most frequently provided non-financial resource was analyzed data or reports, with 22 of 24 programs transmitting this information, although report type, frequency, level of aggregation, and delivery mode varied (e.g., some, but not all, payers developed or provided access to web-based portals for on-demand report creation). The least frequently provided resources were in the domain of care management strategy. Seven programs offered care management personnel or tools—all seven of these programs were primary care-based, and most adhered to the patient-centered medical home model. Programs sponsored by commercial payers tended to have more non-financial supports than those sponsored by governmental insurers.

In terms of the characteristics that led to nine VBP programs being designated as higher- and 15 categorized as lower-intensity (Exhibits 2 and 3, Appendix-Table IIB–b)¹⁷, 17 programs operationalized the spending reduction part of the program as upside-only or two-sided risk, 17 introduced gates and ladders as quality improvement incentives, and 13 introduced four or more non-financial features.

When examining the balance of negative, mixed-negative, null, and mixed-positive outcomes for VBP programs (Exhibit 4), the most frequent observations were mixed-positive (*blue*) or positive (*navy*). For example, of the 14 programs that had quality process measures assessed, 7 were found to have a mixed-positive or positive effect (the sum of 3 mixed-positive and 4 positive effects). Of the 17 programs that had quality utilization measures assessed, 11 were found to have mixed-positive or positive effects. Similarly, in the stratum of the 19 programs which had spending reduction assessed, 14 were found to have mixed-positive or positive effects. When comparing all programs to the stratum involving only VBP programs with rigor > 14, the total number of VBP programs dropped (by 3 each for quality process measures and quality utilization measures and by 4 for spending reduction measures), but that did not shift the overall pattern in which the most frequent program-level finding was mixed-positive or positive.

A slightly different pattern emerged when comparing programs in the higher-intensity stratum to that in the lower-intensity one. First, fewer programs met the criteria for the higher-intensity stratum than the lower-intensity one—for quality process measures, 6 higher-intensity programs were being compared against 8 lower-intensity ones; for quality utilization measures, 8 higher-intensity programs were being compared against 9 lower-intensity ones; and for spending reduction measures, 8 higher-intensity programs were being compared against 14 lower-intensity ones. In these comparisons, higher-intensity programs had mixed-positive or positive outcomes more frequently than did lower-intensity programs—for quality process measures (4 of 6 for high-intensity versus 3 of 8 for low-intensity), for quality utilization measures (6 of 8 for high-intensity versus 5 of 9 for low-intensity), and spending reduction measures (7 of 8 for high-intensity versus 6 of 14 for low-intensity). One lower-intensity VBP program (Maryland Total Patient Revenue) was associated with mixed-negative results for spending measures (i.e., spending increased).

Six VBP programs were evaluated for their effect on patient experience and seven for clinical outcomes—too few in each case to make meaningful comparisons between higher- and lower-intensity programs (Appendix-Section IIIC).¹⁷

In summary, our program-level analysis highlights a generally positive trend for both quality and spending outcomes. Then, when comparing the effects by program intensity, higher-intensity programs were more frequently associated with desired results than lower-intensity ones, and lower-intensity VBP programs were more frequently associated with null effects.

DISCUSSION

To our knowledge, no previous systematic review has examined the variety of spending and quality incentives and non-financial supports introduced in VBP programs and assessed the relationship of program design intensity with desired quality and spending outcomes.

The first finding from this study is the ubiquity of non-financial supports across VBP programs. These supports were present in all but one of the examined programs, indicating that VBP program designers believe that non-financial supports are important alongside financial and quality incentives. Program designers may want to continue to pay attention to the setting in which the program operates and tailor the types of resources provided accordingly. Further research on the particular types or mix of non-financial supports that are most often associated with program success is warranted. We also found that commercial payers generally offered a higher number of non-financial supports than their government counterparts. VBP sponsors in government may want to place additional weight on non-financial supports when designing programs.

Second, this study links VBP with desired outcomes for both quality improvement and spending reduction, not with spending reduction alone as per prior systematic reviews.^{6–16} Third, this study associates higher-intensity VBP programs with mixed-positive or positive effects and lower-intensity VBP programs with null ones.

Taken together, these findings suggest that systematic reviews that take VBP program design intensity into account may yield different take-home messages than those that remain agnostic to overall design. Strengthening the research base for VBP program design may therefore be a worthwhile investment.

Ultimately, stakeholders would like to know if there are specific spending and quality incentives, non-financial supports, or combinations of program elements that would more reliably deliver desired outcomes. Like others, this study just begins to get at that question. Future studies would need to involve a prospective or longitudinal primary data collection with a purposive sample from on-the-ground participants who could report on their organizational contexts and perceptions of both incentives and supports, and a sample large enough that questions about necessary, synergistic or additive characteristics of programs could be answered analytically. Perhaps, as the field grows, this study approach could be repeated in a larger sample of programs.

Nonetheless, this study has implications for payers and delivery systems—higher-intensity programs may affect providers' motivation and ability to generate savings and quality improvements differently than lower-intensity ones. While there may be reasons for payers and providers to introduce different levels of intensity in their VBP programs (e.g., lower-intensity programs may increase voluntary participation; higher-intensity programs may be

harder to implement and may yield unintended negative consequences), these design choices apparently have consequences for spending and quality outcomes.

VBP is a durable policy trend despite mixed results and somewhat halting progress toward its goals. It is sobering to observe that despite two decades of natural experiments, to date, only 24 VBP programs have been empirically evaluated. VBP program design intensity is one aspect of VBP that payers and organizations can control. Going forward, much more can be done to illuminate the features of VBP program design for policymakers and health care providers and to understand when and how they can yield reliable effects in terms of desired care quality and spending.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS:

We acknowledge Dr. Sam Ho, former Chief Medical Officer and Executive Vice President of UnitedHealth Group, Inc. for commissioning the initial review and providing feedback on preliminary findings.

FUNDING:

UnitedHealthcare funded the initial systematic review and characterization of VBP programs. Dr. Chien's Value-Based Payment research program at Harvard Medical School / Boston Children's Hospital funded the refinement of the systematic reviews, characterization of VBP programs, and the examination of the relationship between VBP program design and incentivized quality and spending outcomes.

REFERENCES

1. Rosenthal M, Frank R, Li Z, Epstein A. Early experience with pay-for-performance: from concept to practice. *JAMA*. 2005;294(14):1788–93. [PubMed: 16219882]
2. Song Z, Safran DG, Landon BE, He Y, Ellis RP, Mechanic RE, et al. Health care spending and quality in year 1 of the alternative quality contract. *New England Journal of Medicine*. 2011;365(10):909–18. [PubMed: 21751900]
3. Chien AT, Rosenthal MB. A 3D model for value-based care: the next frontier in financial incentives and relationship support. [Internet]. UnitedHealth Group. 2020 [cited 2022 Nov 1]. Available from: https://newsroom.uhc.com/content/dam/newsroom/Harvard%20Report_FINAL_0923.pdf
4. Lemak CH, Nahra T, Cohen GR, Erb ND, Paustian ML, Share D, et al. Michigan's fee-for-value physician incentive program reduces spending and improves quality in primary care. *Health Aff*. 2015;34(4):645–52.
5. Colla CH, Wennberg DE, Meara E, Skinner JS, Gottlieb D, Lewis VA, et al. Spending differences associated with the Medicare Physician Group Practice Demonstration. *JAMA*. 2012;308(10):1015–23. [PubMed: 22968890]
6. 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 behaviors and patient outcomes (Review). 2011;(7):CD009255.
7. Eijkenaar F, Emmert M, Scheppach M, Schöffski O. Effects of pay for performance in health care: a systematic review of systematic reviews. *Health Policy*. 2013;110(2–3):115–30. [PubMed: 23380190]
8. Damberg CL, Sorbero ME, Lovejoy S, Martsolf G, Raaen L, Mandel D. Measuring Success in Health Care Value-Based Purchasing Programs. *Information Technology and Libraries*. 2014;4(3):60–2.

9. Cattel D, Eijkenaar F. Value-Based Provider Payment Initiatives Combining Global Payments With Explicit Quality Incentives: A Systematic Review. *Medical Care Research and Review*. 2020;77(6):511–37. [PubMed: 31216945]
10. Wilson M, Guta A, Waddell K, Lavis J, Reid R, Evans C. The impacts of accountable care organizations on patient experience, health outcomes and costs: a rapid review. *Journal of Health Services Research and Policy*. 2020;25(2):130–8. [PubMed: 32321282]
11. Yee CA, Pizer SD, Frakt A. Medicare's Bundled Payment Initiatives for Hospital-Initiated Episodes: Evidence and Evolution. *Milbank Quarterly*. 2020;98(3):908–74. [PubMed: 32820837]
12. Vlaanderen FP, Tanke MA, Bloem BR, Faber MJ, Eijkenaar F, Schut FT, et al. Design and effects of outcome-based payment models in healthcare: a systematic review. *European Journal of Health Economics*. 2019;20(2):217–32.
13. Hong YR, 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(8):734–43. [PubMed: 32692140]
14. Mathes T, Pieper D, Morche J, Polus S, Jaschinski T, Eikermann M. Pay for performance for hospitals. *Cochrane Database of Systematic Reviews*. 2019;2019(7).
15. Milad MA, Murray RC, Navathe AS, Ryan AM. Value-Based Payment Models In The Commercial Insurance Sector: A Systematic Review. *Health Affairs*. 2022;41(4):540–8. [PubMed: 35377757]
16. Kaufman BG, Spivack BS, Stearns SC, Song PH, O'Brien EC. Impact of Accountable Care Organizations on Utilization, Care, and Outcomes: A Systematic Review. *Medical Care Research and Review*. 2019;76(3):255–90. [PubMed: 29231131]
17. To access the appendix, click on the Details tab of the article online.
18. Share DA, Mason MH. Michigan's Physician Group Incentive Program offers a regional model for incremental fee for value payment reform. *Health Affairs*. 2012;31(9):1993–2001. [PubMed: 22949448]
19. Rajkumar R, Patel A, Murphy K, Colmers JM, Blum JD, Conway PH, Sharfstein JM. Maryland's All-Payer Approach to Delivery System Reform. *New England Journal of Medicine*. 2014;370(6):493–5. [PubMed: 24410022]
20. Peele P, Keyser D, Lovelace J, Moss D. Advancing Value-Based Population Health Management Through Payer-Provider Partnerships. *Journal for Healthcare Quality*. 2017;40(2):26–32.
21. Dale SB, Ghosh A, Peikes DN, Day TJ, Yoon FB, Taylor EF, et al. Two-Year Costs and Quality in the Comprehensive Primary Care Initiative. *New England Journal of Medicine*. 2016;374(24):2345–56. [PubMed: 27074035]
22. Peikes BD, Dale S, Ghosh A, Taylor EF, Swankoski K, Malley ASO, et al. The Comprehensive Primary Care Initiative: Effects On Spending, Quality, Patients, And Physicians. *Health Affairs*. 2018;37(6):890–99. [PubMed: 29791190]
23. Cuellar A, Helmchen LA, Gimm G, Want J, Burla S, Kells BJ, et al. The CareFirst Patient-Centered Medical Home Program: Cost and Utilization Effects in Its First Three Years. *Journal of General Internal Medicine*. 2016;31(11):1382–8. [PubMed: 27473005]
24. McWilliams JM. Cost Containment and the Tale of Care Coordination. *New England Journal of Medicine*. 2016;375(23):2218–20. [PubMed: 27959672]
25. McConnell KJ. Oregon's Medicaid Coordinated Care Organizations. *JAMA*. 2016; 315(9):869–70 [PubMed: 26847402]
26. Prendergast C The provision of incentives in firms. *Journal of Economic Literature*. 1999;37(1):7–63.
27. Chien AT, Li Z, Rosenthal MB. Improving timely childhood immunizations through pay for performance in medicaid-managed care. *Health Services Resesarch*. 2010;45(6):1934–47.
28. Afendulis CC, Hatfield LA, Landon BE, Gruber J, Landrum MB, Mechanic RE, et al. Early impact of carefirst's patient-centered medical home with strong financial incentives. *Health Affairs*. 2017;36(3):468–75. [PubMed: 28264948]
29. Peele P, Keyser D, Lovelace J, Moss D. Advancing Value-Based Population Health Management Through Payer-Provider Partnerships: Improving Outcomes for Children with Complex Conditions. *Journal for Healthcare Quality*. 2018;40(2):26–32.

30. Salmon RB, Sanderson MI, Walters BA, Kennedy K, Flores RC, Muney AM. A collaborative accountable care model in three practices showed promising early results on costs and quality of care. *Health Affairs*. 2012;31(11):2379–87. [PubMed: 23129667]
31. Maeng DD, Graf TR, Davis DE, Tomcavage J, Bloom FJ. Can a Patient-Centered Medical Home Lead to Better Patient Outcomes? The Quality Implications of Geisinger's ProvenHealth Navigator. *American Journal of Medical Quality*. 2012;27(3):210–6. [PubMed: 21852292]
32. Kuntz G, Tozer JM, Snegosky J, Fox J, Neumann K. Michigan Oncology Medical Home Demonstration Project: First-Year Results. *JCO Oncology Practice*. 2014;10(5):294–7.
33. Dale SB, Ghosh A, Peikes DN, Day TJ, Yoon FB, Taylor EF, et al. Two-Year Costs and Quality in the Comprehensive Primary Care Initiative. *New England Journal of Medicine*. 2016;374(24):2345–56. [PubMed: 27074035]
34. Friedberg MW, Chen PG, White C, Jung O, Raaen L, Hirshman S, et al. Effects of health care payment models on physician practice in the United States. *RAND Health Quarterly*. 2015;5(1):8.
35. 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. *Journal of Epidemiology and Community Health*. 1998;52(6):377–84. [PubMed: 9764259]
36. Weissman JS, Bailit M, D'Andrea G, Rosenthal MB. The design and application of shared savings programs: Lessons from early adopters. *Health Affairs*. 2012;31(9):1959–68. [PubMed: 22949444]



Exhibit 1.

Value-Based Purchasing (VBP) number of unique empirical evaluations and type of outcomes studied, by higher- and lower-intensity programs

Source: Authors' summary of our systematic review inclusions used to inform this paper (methodology and citations of included programs and manuscripts available in appendix).¹⁷

Note: Unique empirical evaluations sum to less than the total number of evaluations for spending and quality outcomes as a given study may evaluate both spending and quality. The median number of evaluations per program was 2. The mean and standard deviation were 3±2.3. The greatest number of evaluations for a single program was 10.

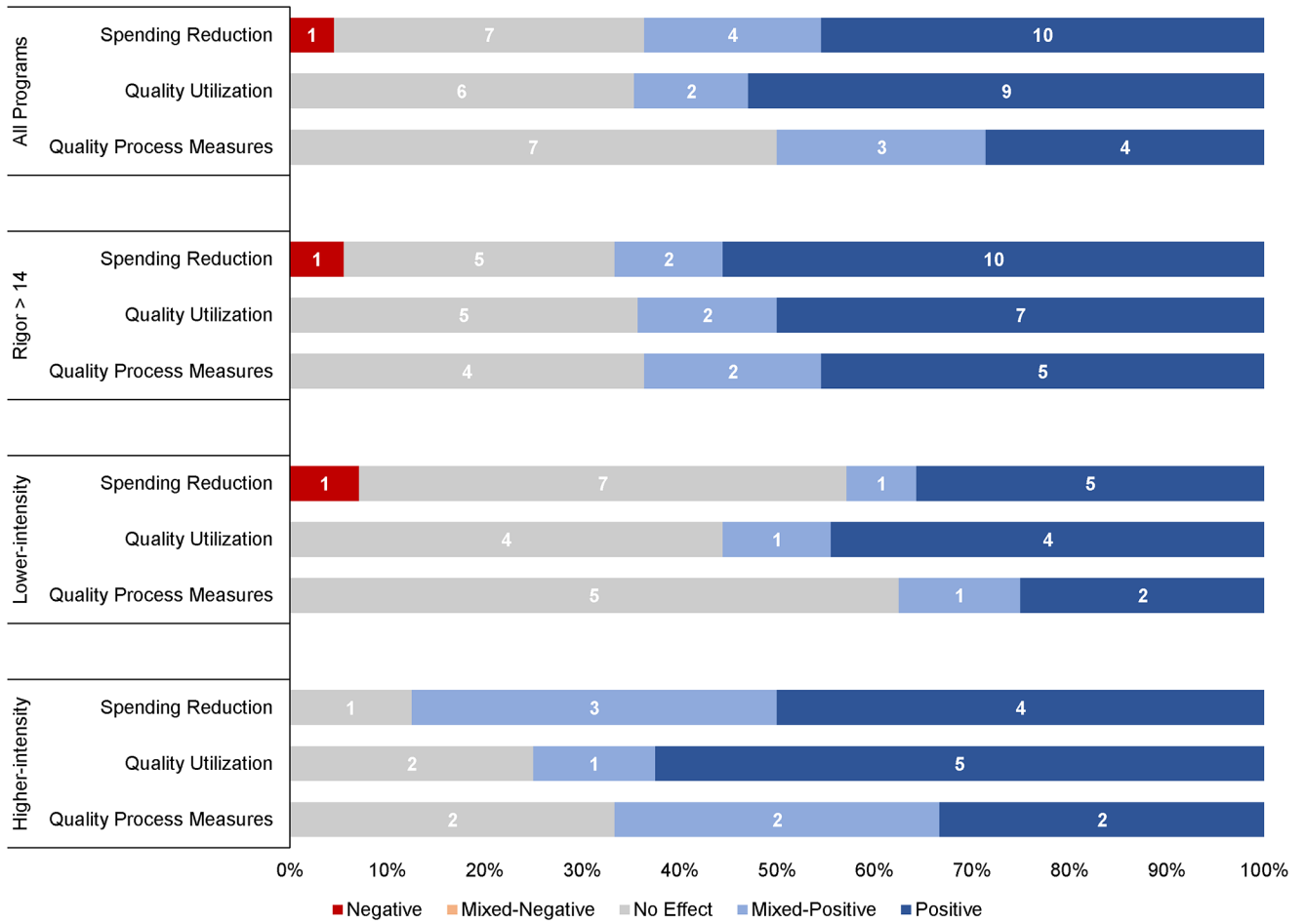


Exhibit 4. Value-Based Purchasing (VBP) outcome positivity by literature rigor and program intensity
Source: Authors’ summary of qualitative analyses conducted after assessing each included manuscript for outcomes, significance, and directionality
Note: The categories “All,” “Higher-intensity,” and “Lower-intensity” include all studies irrespective of rigor level

Value-Based Purchasing (VBP) spending and quality incentives, by higher- versus lower-intensity program design

Exhibit 2:

	Spending Incentives					Quality Incentives			
	Resource Efficiency Incentive	Prospective Payment System	Bunded Payment	Upside-only Risk	Two-sided Risk	Pay for Performance	Gate	Ladder	
Higher Intensity Programs									
Blue Cross Blue Shield Alternative Quality Contract	NO	NO	NO	NO	YES	YES	NO	YES	
Maryland Multipayer Patient Centered Medical Home	NO	NO	NO	NO	YES	NO	NO	YES	
Capital District Physicians' Health Plan	NO	NO	NO	YES	YES	NO	NO	YES	
Oregon Medicaid Coordinated Care Organizations	NO	NO	NO	NO	YES	NO	NO	YES	
Medicare Shared Savings Program	NO	NO	NO	YES	YES	NO	YES	YES	
Blue Cross Blue Shield CareFirst Total Care and Cost Improvement	NO	NO	NO	YES	NO	YES	YES	YES	
Cigna Collaborative Accountable Care	NO	NO	NO	YES	NO	NO	YES	YES	
Geisinger ProvenHealth Navigator	NO	NO	NO	YES	NO	YES	YES	YES	
Pennsylvania Chronic Care Initiative	NO	NO	NO	YES	NO	NO	YES	YES	
Lower Intensity Programs									
Maryland Total Patient Revenue	NO	NO	NO	NO	YES	YES	NO	NO	
Medicare Pioneer Accountable Care Organizations Model	NO	NO	NO	NO	YES	NO	YES	YES	
Maryland Global Hospital	NO	NO	NO	NO	YES	YES	NO	NO	
Comprehensive Primary Care Initiative	NO	NO	NO	YES	NO	NO	YES	NO	
Michigan Oncology Medical Home	NO	NO	NO	YES	NO	NO	YES	YES	
Maryland Quality-Based Reimbursement Program	NO	NO	NO	YES	NO	YES	NO	NO	
Nursing Home Value-Based Purchasing	NO	NO	NO	YES	NO	NO	YES	YES	
Physician Group Practice	NO	NO	NO	YES	NO	NO	NO	YES	
Michigan Physician Group Incentive	YES	NO	NO	NO	NO	YES	NO	NO	
Comprehensive Care for Joint Replacement	NO	NO	YES	NO	NO	NO	NO	YES	

	Spending Incentives					Quality Incentives			
	Resource Efficiency Incentive	Prospective Payment System	Bundled Payment	Upside-only Risk	Two-sided Risk	Pay for Performance	Gate	Ladder	
Allegheny High-Value Care for Kids	NO	NO	YES	NO	NO	NO	YES	NO	
Total Joint Arthroplasty Bundled Payment Pilot	NO	NO	YES	NO	NO	NO	NO	YES	
Hospital Readmission Reduction Program	NO	YES	NO	NO	NO	YES	NO	NO	
Hospital Value-Based Purchasing	YES	YES	NO	NO	NO	YES	NO	NO	
Physician Value-Based Modifier Program	YES	YES	NO	NO	NO	YES	NO	NO	
<i>sum</i>	3	3	3	11	8	10	11	14	

Source: Author's summary and characterization of spending and quality incentives of each included program found via literature search (citations included in appendix).¹⁷

Note: A total of 17 programs had upside-only/two-sided risk and a total of 17 programs had quality incentives tied to gates or ladders.

Exhibit 3:

Value-Based Purchasing (VBP) non-financial supports, by higher- versus lower-intensity program design

	Non-financial Supports					
	Analyzed Data	Technical Assistance	Infrastructure Payments	Raw Data	Risk Management Support	Care Management
Higher-Intensity Programs						
Blue Cross Blue Shield Alternative Quality Contract	YES	YES	YES	YES	YES	NO
Maryland Multipayer Patient Centered Medical Home	YES	YES	YES	NO	YES	YES
Capital District Physicians' Health Plan	YES	YES	YES	YES	NO	YES
Oregon Medicaid Coordinated Care Organizations	YES	YES	YES	YES	NO	NO
Medicare Shared Savings Program	YES	YES	NO	YES	YES	NO
Blue Cross Blue Shield CareFirst Total Care and Cost Improvement	YES	YES	YES	YES	YES	YES
Cigna Collaborative Accountable Care	YES	YES	YES	YES	YES	YES
Geisinger ProvenHealth Navigator	YES	YES	YES	NO	YES	YES
Pennsylvania Chronic Care Initiative	YES	YES	YES	YES	NO	NO
Lower-Intensity Programs						
Maryland Total Patient Revenue	YES	YES	NO	YES	NO	NO
Medicare Pioneer Accountable Care Organizations Model	NO	YES	NO	YES	YES	NO
Maryland Global Hospital	YES	YES	NO	YES	NO	NO
Comprehensive Primary Care Initiative	YES	YES	YES	YES	NO	NO
Michigan Oncology Medical Home	YES	NO	YES	NO	NO	YES
Maryland Quality-Based Reimbursement Program	YES	YES	NO	YES	NO	NO
Nursing Home Value-Based Purchasing	YES	YES	NO	NO	NO	NO
Physician Group Practice	YES	NO	NO	YES	NO	NO
Michigan Physician Group Incentive	YES	YES	YES	YES	NO	NO
Comprehensive Care for Joint Replacement	YES	YES	NO	YES	YES	NO
Allegheny High-Value Care for Kids	YES	YES	YES	NO	NO	YES
Total Joint Arthroplasty Bundled Payment Pilot*	NO	NO	NO	NO	NO	NO
Hospital Readmission Reduction Program	YES	YES	NO	NO	NO	NO

Non-financial Supports						
	Analyzed Data	Technical Assistance	Infrastructure Payments	Raw Data	Risk Management Support	Care Management
Hospital Value-Based Purchasing	YES	YES	NO	NO	NO	NO
Physician Value-Based Modifier Program	YES	NO	NO	NO	NO	NO
<i>sum</i>	22	20	12	15	8	7

Source: Authors' counts of non-financial program supports discovered via literature search (citations included in appendix)¹⁷

Note:

* The Total Joint Arthroplasty Bundled Payment Pilot was the only VBP program for which we could not locate descriptions of non-financial supports via publicly available sources. A total of 13 programs had 4 or more non-financial supports.