



Published in final edited form as:

Am J Manag Care. 2017 January ; 23(1): 48–53.

Value-based payment in implementing evidence-based care: the Mental Health Integration Program in Washington State

Yuhua Bao^{1,2}, Thomas G. McGuire³, Ya-Fen Chan⁴, Ashley A. Eggman¹, Andrew M. Ryan⁵, Martha L. Bruce², Harold Alan Pincus^{6,7}, Erin Hafer⁸, and Jürgen Unützer⁴

¹Department of Healthcare Policy and Research, Weill Cornell Medical College, New York, New York

²Department of Psychiatry, Weill Cornell Medical College, New York, New York

³Department of Health Care Policy, Harvard Medical School, Boston, Massachusetts

⁴Department of Psychiatry and Behavioral Science, University of Washington School of Medicine, Seattle, Washington

⁵Department of Health Management and Policy, School of Public Health, University of Michigan, Ann Arbor, Michigan

⁶Department of Psychiatry, Columbia University College of Physicians and Surgeons, New York, New York

⁷NewYork-Presbyterian Hospital, New York, New York

⁸Community Health Plan of Washington, Seattle, Washington

Abstract

Objectives—To assess the role of VBP in improving fidelity and patient outcomes in community implementation of the Collaborative Care Model (CCM), an evidence-based mental health intervention.

Study Design—Retrospective study based on a natural experiment

Methods—We used clinical tracking data of 1,806 adult patients enrolled in a large implementation of CCM in community health clinics in Washington State. VBP was initiated in Year 2 of the program, creating a natural experiment. We compared implementation fidelity (measured by three process-of-care elements of the CCM) between patient-months exposed to VBP and patient-months not exposed to VBP. A series of regressions were estimated to check robustness of findings. We estimated a Cox proportional hazard model to assess the effect of VBP on time to achieving clinically significant improvement in depression (measured based on changes in depression symptom scores over time).

Results—Estimated marginal effects of VBP on fidelity ranged from 9% to 30% of the level of fidelity had there been no exposure to VBP ($p < 0.05$ for every fidelity measure). Improvement in fidelity in response to VBP was greater among providers with a larger patient panel and among

*Corresponding author: Yuhua Bao, PhD. 402 E 67th St., New York, NY 10065; (O) 646 962-8037; yub2003@med.cornell.edu.

providers with a lower level of fidelity at baseline. Exposure to VBP was associated with an adjusted hazard ratio of 1.45 (95% CI: 1.04–2.03) for achieving clinically significant improvement in depression.

Conclusions—VBP improved fidelity to key elements of the CCM, both directly and not directly incentivized by the VBP, and improved patient depression outcomes.

Despite extraordinary increases in medical knowledge, health care in the United States frequently falls short of evidence-based standards.¹ Implementation fidelity is defined as the degree to which interventions or programs are implemented as intended by the program developers.² Poor implementation fidelity is one explanation why the promise of evidence-based medicine remains unfulfilled.^{3,4} Substantial variation among providers exists in the intensity of implementation and degree of fidelity to the evidence-base.^{5–8} Current financial incentives in U.S. health care system contribute to a poor “business case” to adopt evidence-based practices in an effective manner.⁹ Targeted financial incentives have the potential to improve fidelity and improve implementation effectiveness.

Value-based payment (VBP), a form of pay-for-performance, incentivizes quality and outcomes of care by tying payment to providers with pre-defined quality or efficiency targets. VBP has been widely adopted by private and public payers. Notable recent examples include Medicare’s launch of the hospital VBP program in 2013¹⁰ and the announcement by the Secretary of Health and Human Services of measurable goals and a timeline to move the health care system at large towards quality-based payment.¹¹ Majority of existing VBP programs provide stand-alone financial incentives without a support system to providers for care process redesign. VBPs embedded in implementation of evidence-based care and designed to improve implementation effectiveness are rare^{12–15} and have not been extensively studied.

In this study, we assessed whether a value-based payment component could improve the effectiveness of implementation of the Collaborative Care Model (CCM). The CCM is a team-based approach to treating depression and other common behavioral health conditions in primary care and has a strong evidence base.^{16–18} The CCM team includes the primary care physician, a care manager, and a consulting psychiatrist. The key principles of the model include systematic follow-up of the patients by the care manager, measurement-based care that uses symptom rating scales to track clinical improvements or identify patients not improving,¹⁹ and “stepped care”²⁰ with which treatment is systematically adjusted or intensified (by the primary care team but with input from the consulting psychiatrist) for patients not improving.²¹ Implementation of CCM has gained momentum in recent years. In its proposed rules for the 2016 Physician Fee Schedule, the Centers for Medicare and Medicaid Services stated an intention to modify current payment to cover the CCM.²²

We conducted our assessment in the context of the Mental Health Integration Program (MHIP) of Washington State.²³ The MHIP is an ongoing, publicly funded implementation of CCM in a diverse network of community health clinics across the state of Washington. Initially started in 2008 in the two most urban counties in the state, MHIP has now expanded state-wide and served a total of over 35,000 individuals. The VBP component of the MHIP payment started in 2009 (second year of MHIP) in response to the substantial variation in

quality of care and patient outcomes seen in the first year. It adopted some best practice design features of value-based payment^{24–26} including close mapping of quality measures with key elements of the evidence-based model, a substantial size of the incentive payments (25% of total payments to providers), and a dynamic set of quality measures and targets that are adjusted over time. It thus offers a unique opportunity to assess the role of VBP in improving implementation effectiveness of evidence-based care in community settings.

We hypothesized that MHIP VBP improved fidelity to CCM measured by important process-of-care elements of the model, both incentivized and not directly incentivized by the VBP. We further hypothesized that MHIP VBP improved patient depression outcomes. Provider organizations with a larger patient panel have more at stake under a VBP scheme as well as greater resources to invest in quality improvement in response to VBP.²⁷ Providers with a lower level of performance at baseline have more room and may be more motivated to improve.^{28–30} We thus hypothesized that the effect of VBP on implementation fidelity was greater among clinics with a larger MHIP patient caseload and among clinics with a lower level of fidelity prior to VBP.

Methods

Study Period, Population, and Data

This study focuses on Phase 1 of the MHIP VBP. Phase 1 of the program (covering Year 2009) used 4 process-of-care targets mapping closely with the principles of the CCM (Table 1). Provider organizations would initially receive 75% of their total payment for CCM (i.e., with 25% hold-back) and receive 5% for achieving each of the 4 targets in a calendar quarter with an additional 5% being awarded for participation. Adjustments to the VBP scheme were made in subsequent phases by raising benchmarks for existing targets, eliminating targets that had been achieved by most provider organizations, and/or adding new targets to address emerging gaps in quality, thus providing incentives for continuous improvement.

Our study population were patients 18 years or older who initiated care in MHIP between January 1, 2008 and June 30, 2009 in one of the 35 community health clinics that started MHIP implementation in 2008 (and therefore had experience with MHIP both before and after the launch of Phase 1 VBP). We restricted patient enrollment in MHIP to June 30, 2009 to ensure that the first 6 months of care for every patient enrolled in 2009 were under the influence of Phase 1 (and not Phase 2) of VBP. All 35 clinics were located in King and Pierce counties, the two most populous counties in Washington State, and were affiliated with 7 community health centers. These community health centers were the parent organizations of the clinics and were Federally Qualified Health Centers. The population they served were primarily patients with Medicaid, other state funded programs, and patients who were uninsured. Clinical social workers, psychologists, licensed mental health counselors, and other clinicians staffed at the clinics served as the CCM care managers in MHIP. Patient inclusion criteria included a baseline Patient Health Questionnaire (or PHQ-9³¹) score of 10 or greater indicating clinically significant depression and at least one follow-up contact with the MHIP care manager within 6 months/24 weeks of the initial contact (to allow at least one chance to assess depression outcomes). Patients whose last contact with MHIP occurred within one week from the first contact were further excluded

since they likely were determined ineligible for MHIP. The vast majority of the MHIP patients during our study period were enrollees in Washington State's Disability Lifeline Program. These patients were temporarily disabled because of a physical or mental health condition and expected to be unemployed for 90 days or more. King County extended eligibility to additional patient populations including low-income mothers and their children, low-income older adults, uninsured, veterans, and veterans' family members.

Our data were from the web-based registry³² used by all MHIP participating clinics to systematically document care management activities and clinical outcomes and to assist with population management.

Measures

Three dichotomous measures at the patient-month level captured fidelity to major domains of the CCM (Table 1). "At least one follow-up contact with the care manager" reflects the principle of systematic follow-up; "at least one psychiatric consultation" reflects the principle of "stepped care" – the idea that treatment should be systematically changed or intensified for patients not responding to initial treatment; an important mechanism by which stepped care is operationalized in CCM is through consultation with a mental health specialist (usually a psychiatrist) for potential treatment changes. These two measures were closely related to the two quality targets in Phase 1 of MHIP VBP (Table 1). "At least one PHQ-9 assessment" reflects the principle of measurement-based care, whereby treatment teams use symptom rating scales to systematically track clinical improvements or lack thereof. This measure was not explicitly incentivized in MHIP VBP. Data on current medications (also documented in the MHIP registry) are not available for research at this point. We therefore had no fidelity measure mapping the fourth VBP measure, documentation of current psychiatric medication in registry for 75% of cases (Table 1). Each fidelity measure was assessed for each 4-week interval starting from the patient's initial contact with the MHIP care manager up to 24 weeks or until the patient's last contact with the care manager, whichever occurred first.

We tested hypotheses about two potential modifiers of the effects of VBP on fidelity: size of the MHIP patient panel, measured by cumulative number of patients treated at the clinic prior to Phase 1 VBP (i.e., in 2008) and clinic-level fidelity at baseline, measured by the average count of follow-up contacts, psychiatric consultation, or PHQ assessments over the first months of all patients treated at the clinic in 2008.

A clinically significant improvement in depression was defined as achieving a follow-up PHQ-9 score of less than 10 or achieving a 50% or more reduction in PHQ-9³¹ within 24 weeks of the initial care manager contact ("improvement" hereafter).

Statistical Analysis

The rolling enrollment of patients in the MHIP program throughout our study period created a natural experiment. We compared the fidelity outcomes for patient-months exposed to VBP with patient-months not exposed to VBP. We estimated a multi-level linear probability model for each fidelity outcome with a random intercept at the patient level to account for clustering of months of the same patient. The key independent variable is a dichotomous

indicator of VBP exposure, defined as 1 if the index patient-month started after 1/1/2009, when Phase 1 VBP took effect, and 0 otherwise. Because care management activities were more intensive in early months than later months of a patient's treatment episode,³³ our adjusted analysis controlled for a set of dummy variables indicating whether the index month was the patient's 1st, 2nd, ..., or, 6th month in MHIP. We conducted several sensitivity analyses. We restricted the sample to Disability Lifeline enrollees (accounting for 85% of the entire sample). We also estimated the linear probability model with a fixed effect for each patient, first with the entire sample (i.e., regardless of whether a patient had exposure to VBP), then restricting to patients who contributed months both pre- and post- VBP during the first 24 weeks of their care. The latter approach allowed us to examine the effects of VBP on treatment fidelity within the same patient but was conducted with a much smaller sample size (about 1/3 of the original sample) and thus subject to lower precision in estimates. We also estimated the logistic version of each model and compared implications (e.g., marginal effect of VBP) based on both sets of analyses.

Hypotheses regarding the modifying effects were tested by adding interaction terms between the VBP indicator and the hypothesized modifier to the models. We included an additional interaction between VBP and the quadratic term of the modifier (e.g., cumulative caseload squared) to allow for nonlinear effects. The VBP incentives were directed at the community health centers (also known as "implementation sites" in MHIP; 7 in total in our study sample); each site had multiple clinics. We conducted sensitivity analysis where hypothesized modifiers (caseload and baseline fidelity) were measured at the site level.

To assess the association between VBP and depression outcomes, we estimated an extended Cox proportional hazard model of time to improvement in depression, censored at 24 weeks after the initial assessment/contact or patient's last contact with the MHIP care manager, whichever occurred first. The key independent variable was an indicator of VBP exposure (i.e., 1 after 01/01/2009, 0 otherwise). This indicator varied over time (switched from 0 to 1) for patients who enrolled in MHIP in 2008 but whose observation period ended in 2009. We conducted tests of the proportional hazard assumption based on the Schoenfeld residuals.³⁴

Adjusted analyses controlled for baseline patient age and gender, baseline PHQ-9 scores and comorbid behavioral health conditions, MHIP eligibility categories, and clinic fixed effects (to control for between-clinic differences in quality of care). To control for possible clinic learning over time, we also included the number of months the clinic had been participating in MHIP at the time the index patient was enrolled in MHIP and its quadratic and cubic terms.

Results

Table 2 presents descriptive statistics of baseline patient characteristics for the entire sample (n=1806) and by whether a patient had at least one month of exposure to VBP within 24 weeks since their first contact with MHIP care managers. Patients with no exposure to VBP were more likely to be enrollees in the Disability Lifeline program than patients with at least one month of exposure (96.2% vs. 79.7%). Partly because of this difference, patients with no exposure were more likely to be between 40–59 years old and more likely to have a PHQ-9

score of 20 or plus (indicating severe depression symptoms), compared to patients without any exposure. Prevalence of comorbid behavioral conditions was comparable between the two cohorts except that rates of anxiety and bipolar disorders were slightly lower among the no-exposure group than the exposed group.

For the fidelity outcomes, results of multivariate analyses were presented in the form of predicted probabilities with and without exposure to VBP and the marginal effect of VBP (Table 3). (The mixed-effects linear probability models and their logistic counterparts generated very similar results; results reported hereafter and in Table 3 were based on the linear probability models.) Based on analyses conducted with the entire sample, the effect of VBP on the probability of at least one follow-up contact, psychiatric consultation, and PHQ assessment in a month was an increase of 0.05 (95% confidence interval or CI: 0.00–0.10; $p < 0.05$), 0.04 (95% CI: 0.00–0.07; $p < 0.05$), and 0.07 (95% CI: 0.02–0.11; $p < 0.05$), respectively. The magnitude of the increase was about 9%, 30%, and 15% of the respective fidelity outcome had there been no exposure to VBP. Analysis restricting to Disability Lifeline enrollees (85% of the unrestricted sample) generated very similar results, with slightly smaller marginal effects of VBP for follow-up contacts and PHQ assessments, but slightly greater marginal effect for psychiatric consultation (Table 3).

Sensitivity analysis of the fidelity outcomes by controlling for patient fixed effects produced similar or slightly greater point estimates of the marginal effects of VBP (eAppendix Table). However, fixed-effects analysis restricting to patients who received care both before and after VBP had a much reduced sample size (359 patients compared to 1,806 in the unrestricted fixed effects analysis). Marginal effects of VBP were not statistically significant for follow-up contacts or psychiatric consultation, but remained strong for PHQ assessments (0.09, 95% CI: 0.02–0.16).

Our analysis indicated that both the size of the MHIP caseload at a clinic and the level of fidelity prior to VBP modified the effect of the VBP. As shown in eAppendix Figure A, for follow-up contacts and PHQ assessments, the marginal effect of VBP increased with the number of patients treated at the clinic prior to VBP. For follow-up contacts, marginal effect of VBP did not achieve statistical significance until the number of patients treated at clinic in 2008 was 100 or more (top 25% of clinics); for PHQ assessments, not until 140+ (top 10% of clinics). Caseload did not seem to modify the VBP effect for psychiatric consultation. On the other hand, for each fidelity measure, the marginal effect of VBP decreased with the level of fidelity at the clinic prior to the start of VBP (eAppendix Figure B). For example, the effect of VBP on follow-up contacts was statistically greater than zero only among clinics whose first-month follow-up contacts in 2008 averaged below 0.8 (accounting for 75% of all clinics). Sensitivity analysis defining modifiers at the implementation site level produced largely consistent findings. One exception was that, for psychiatric consultation, the VBP effect seemed to decrease with the size of MHIP caseload at the site level, whereas there was no clear modifying effect of caseload defined at the clinic level.

Consistent with results for the fidelity outcomes, exposure to VBP was associated with an adjusted hazard ratio (AHR) of 1.45 (95% CI: 1.04–2.03) for achieving clinically significant improvement in depression, indicating that exposure to VBP was associated with a shorter

time to improvement. This result held when we restricted the sample to Disability Lifeline patients (AHR=1.47, 95% CI: 1.03–2.12).

Discussion

With data from a state-wide implementation of the Collaborative Care Model in community health clinics, we found that a VBP program embedded in a community-based implementation improved fidelity to several key process-of-care elements of the evidence-based model, both directly incentivized and not directly incentivized by the VBP. Consistent with our hypotheses, we also found stronger responses to VBP among provider organizations that cared for a larger number of patients and among organizations with a lower level of initial fidelity. Finally, we found that VBP led to better patient outcomes indicated by a shorter time to clinically significant improvement in depressive symptoms.

Our findings of VBP effects contrast with the limited evidence supporting the effectiveness of existing VBP programs.^{15,25,35–37} Several reasons may underlie the differences in findings. First, the MHIP paired financial incentives with chronic care quality improvement and capacity-building efforts. An expert team at the University of Washington provided training to care managers from all participating community health clinics and made archived training materials available online. Consulting psychiatrists were arranged in a contractual relationship to work with all participating clinics. The clinical tracking system is a crucial tool to enable population health management and case tracking. Implementation of the tracking system was a precondition for receiving funding for MHIP and was achieved at all clinics. Existing VBP contracts, on the other hand, typically do not provide a support system for quality improvement. Second, the MHIP VBP targeted several key elements of a single evidence-based care model, focusing improvement efforts and sending strong signals and clear directions to provider organizations regarding what to improve. Existing VBP programs typically contain a large number of quality targets that may not be clinically meaningful, dissipating incentives and failing to engage clinicians.³⁸

Monthly PHQ assessments, the measure not directly incentivized under MHIP VBP, improved by 15% in response to VBP. Since PHQ assessment was conducted at follow-up contacts with the care manager, incentivizing systematic follow-up may have had the “spillover effects” of incentivizing PHQ assessments. A direct implication is that, for fidelity/process-of-care measures that complement one another, designers of VBP programs may consider recognizing some but not all of them as VBP targets. Keeping the target set parsimonious (and thus not diluting incentives) may not need to come at the price of forsaking important quality goals.

Clinics with a larger patient caseload responded to VBP more in two of three fidelity measures considered, suggesting that smaller clinics may perceive insufficient incentives because of the limited scope of their VBP payment²⁵ and/or may lack resources to make systematic changes to care in response to VBP. To ensure that provider organizations of all sizes (and their patients) benefit from VBP, implementation initiatives may consider pooling resources, for example, by establishing learning collaboratives and providing coaching and consultation to organizations in need. Consistent with our hypothesis and findings of

previous studies,^{25,28–30} lower baseline fidelity was associated with greater improvement in fidelity in response to VBP, thus reducing the variation in fidelity/quality among provider organizations implementing the evidence-based model. While a desirable outcome, it also reveals the fact that, with a single performance target for all providers, high-performers may not be adequately motivated to improve further even though there is still room for improvement. One option would be to have two sets of thresholds to be applied to provider organizations with different initial levels of performance.³⁹ This option, however, adds to the complexity of the VBP and may be perceived unfair or unacceptable, especially by high-performing provider organizations.

Our study has limitations. We assessed the effects of VBP in a natural experiment, not a randomized trial. Although we controlled for important differences in patient baseline characteristics, differences among clinics that do not change over time (with clinic fixed effects), and proxies for provider learning over time, these controls were not perfect. However, a series of sensitivity analysis demonstrated the robustness of our findings. The fidelity and patient outcomes we examined were subject to data availability and usability. In particular, we were not able to examine antidepressant management as an important fidelity outcome at this point but intend to do so in the future when data become available. We used data from the implementation of a specific care management model in a single state, potentially limiting the generalizability of our findings to other evidence-based care approaches or other geographic areas. This limitation, however, is mitigated by the fact that the CCM is highly consistent with the Chronic Care Model^{40,41} and that the MHIP involved a large and diverse set of provider organizations.

In conclusion, our study provided strong evidence that a value-based payment component adopting best practices of VBP design and embedded in an implementation initiative was effective in improving fidelity to key elements of the evidence-based model, both directly and not directly incentivized by the VBP, and, in turn, improved patient outcomes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Blumenthal D, McGinnis JM. Measuring Vital Signs: an IOM report on core metrics for health and health care progress. *JAMA*. 2015; 313:1901–2. [PubMed: 25919301]
2. Dusenbury L, Brannigan R, Falco M, Hansen WB. A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health Education Research*. 2003; 18:237–56. [PubMed: 12729182]
3. Chin MH, Cook S, Drum ML, et al. Improving diabetes care in midwest community health centers with the health disparities collaborative. *Diabetes care*. 2004; 27:2–8. [PubMed: 14693957]
4. Landon BE, Hicks LS, O'Malley AJ, et al. Improving the management of chronic disease at community health centers. *N Engl J Med*. 2007; 356:921–34. [PubMed: 17329699]
5. Rubenstein LV, Jackson-Triche M, Unutzer J, et al. Evidence-based care for depression in managed primary care practices. *Health Aff (Millwood)*. 1999; 18:89–105.
6. Lichtman JH, Roumanis SA, Radford MJ, Riedinger MS, Weingarten S, Krumholz HM. Can practice guidelines be transported effectively to different settings? Results from a multicenter

- interventional study. *The Joint Commission journal on quality improvement*. 2001; 27:42–53. [PubMed: 11147239]
7. Pearson M, Wu S, Schaefer J, et al. Assessing the implementation of the chronic care model in quality improvement collaboratives. *Health services research*. 2005; 40:978–96. [PubMed: 16033488]
 8. Chin MH, Auerbach SB, Cook S, et al. Quality of diabetes care in community health centers. *American journal of public health*. 2000; 90:431–4. [PubMed: 10705866]
 9. Leatherman S, Berwick D, Iles D, et al. The business case for quality: case studies and an analysis. *Health Aff (Millwood)*. 2003; 22:17–30. [PubMed: 12674405]
 10. [Accessed January 28, 2016] Hospital Value-Based Purchasing. 2014. at <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/hospital-value-based-purchasing/index.html?redirect=/Hospital-Value-Based-Purchasing/>
 11. U.S. Department of Health and Human Services. Better, Smarter, Healthier: In historic announcement, HHS sets clear goals and timeline for shifting Medicare reimbursements from volume to value. Washington, D.C: 2015.
 12. An LC, Bluhm JH, Foldes SS, et al. A randomized trial of a pay-for-performance program targeting clinician referral to a state tobacco quitline. *Archives of internal medicine*. 2008; 168:1993–9. [PubMed: 18852400]
 13. Beaulieu ND, Horigan DR. Putting smart money to work for quality improvement. *Health services research*. 2005; 40:1318–34. [PubMed: 16174136]
 14. Roski J, Turbyville S, Dunn D, Krushat M, Scholle SH. Resource use and associated care effectiveness results for people with diabetes in managed care organizations. *Am J Med Qual*. 2008; 23:365–74. [PubMed: 18820141]
 15. Scott A, Sivey P, Ait Ouakrim D, et al. The effect of financial incentives on the quality of health care provided by primary care physicians. *Cochrane Database Syst Rev*. 2011:CD008451. [PubMed: 21901722]
 16. Archer J, Bower P, Gilbody S, et al. Collaborative care for depression and anxiety problems. *Cochrane Database Syst Rev*. 2012; 10:CD006525. [PubMed: 23076925]
 17. Gilbody S, Bower P, Fletcher J, Richards D, Sutton AJ. Collaborative care for depression: a cumulative meta-analysis and review of longer-term outcomes. *Archives of internal medicine*. 2006; 166:2314–21. [PubMed: 17130383]
 18. Williams LS, Kroenke K, Bakas T, et al. Care management of poststroke depression: a randomized, controlled trial. *Stroke; a journal of cerebral circulation*. 2007; 38:998–1003.
 19. Trivedi MH. Tools and strategies for ongoing assessment of depression: a measurement-based approach to remission. *The Journal of clinical psychiatry*. 2009; 70(Suppl 6):26–31. [PubMed: 19922741]
 20. Von Korff M, Tiemens B. Individualized stepped care of chronic illness. *The Western journal of medicine*. 2000; 172:133–7. [PubMed: 10693379]
 21. Katon W, Unutzer J, Wells K, Jones L. Collaborative depression care: history, evolution and ways to enhance dissemination and sustainability. *General hospital psychiatry*. 2010; 32:456–64. [PubMed: 20851265]
 22. Federal Register. Department of Health and Human Services and Centers for Medicare & Medicaid Services. Medicare Program; Revisions to Payment Policies Under the Physician Fee Schedule and Other Revisions to Part B for CY 2016. 2015. p. 41685-966.
 23. Washington State Mental Health Integration Program (MHIP). [Accessed January 28, 2016] at <http://integratedcare-nw.org/index.html>
 24. Christianson JB, Leatherman S, Sutherland K. Lessons from evaluations of purchaser pay-for-performance programs: a review of the evidence. *Med Care Res Rev*. 2008; 65:5S–35S. [PubMed: 19015377]
 25. Van Herck P, De Smedt D, Annemans L, Remmen R, Rosenthal MB, Sermeus W. Systematic review: Effects, design choices, and context of pay-for-performance in health care. *BMC Health Serv Res*. 2010; 10:247. [PubMed: 20731816]
 26. de Bruin SR, Baan CA, Struijs JN. Pay-for-performance in disease management: a systematic review of the literature. *BMC Health Serv Res*. 2011; 11:272. [PubMed: 21999234]

27. Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O. Diffusion of innovations in service organizations: systematic review and recommendations. *Milbank Q.* 2004; 82:581–629. [PubMed: 15595944]
28. Greene J, Hibbard JH, Overton V. Large performance incentives had the greatest impact on providers whose quality metrics were lowest at baseline. *Health Aff (Millwood).* 2015; 34:673–80. [PubMed: 25847651]
29. Li J, Hurley J, DeCicca P, Buckley G. Physician response to pay-for-performance: evidence from a natural experiment. *Health Econ.* 2014; 23:962–78. [PubMed: 23861240]
30. Rosenthal MB, Frank RG, Li Z, Epstein AM. Early experience with pay-for-performance: from concept to practice. *JAMA.* 2005; 294:1788–93. [PubMed: 16219882]
31. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *Journal of general internal medicine.* 2001; 16:606–13. [PubMed: 11556941]
32. Unutzer J, Choi Y, Cook IA, Oishi S. A web-based data management system to improve care for depression in a multicenter clinical trial. *Psychiatric services.* 2002; 53:671–3. 8. [PubMed: 12045303]
33. Bao Y, Casalino LP, Ettner SL, Bruce ML, Solberg LI, Unutzer J. Designing payment for Collaborative Care for Depression in primary care. *Health services research.* 2011; 46:1436–51. [PubMed: 21609327]
34. Grambsch PM, Therneau TM. Proportional Hazards Tests and Diagnostics Based on Weighted Residuals. *Biometrika.* 1994; 81:515–26.
35. Jha AK, Joynt KE, Orav EJ, Epstein AM. The long-term effect of premier pay for performance on patient outcomes. *N Engl J Med.* 2012; 366:1606–15. [PubMed: 22455751]
36. 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:CD009255. [PubMed: 21735443]
37. Houle SK, McAlister FA, Jackevicius CA, Chuck AW, Tsuyuki RT. Does performance-based remuneration for individual health care practitioners affect patient care?: a systematic review. *Annals of internal medicine.* 2012; 157:889–99. [PubMed: 23247940]
38. Jha AK. Time to get serious about pay for performance. *JAMA.* 2013; 309:347–8. [PubMed: 23340633]
39. Ryan AM. Will value-based purchasing increase disparities in care? *N Engl J Med.* 2013; 369:2472–4. [PubMed: 24369072]
40. Von Korff M, Gruman J, Schaefer J, Curry SJ, Wagner EH. Collaborative management of chronic illness. *Annals of internal medicine.* 1997; 127:1097–102. [PubMed: 9412313]
41. Wagner EH, Austin BT, Von Korff M. Organizing care for patients with chronic illness. *The Milbank quarterly.* 1996; 74:511–44. [PubMed: 8941260]

Table 1

Quality targets of Phase 1 VBP of the Mental Health Integration Program and corresponding fidelity measures

Domain of Collaborative Care	Phase 1 VBP Target	Fidelity Measure
Systematic follow-up	1. Maintain minimum monthly caseload	At least one follow-up contact with care manager in each 4-week period
	2. 50% caseload receives 2 visits to care manager per month	Same as above
Measurement-based care	Not incentivized	At least one PHQ-9 assessment in each 4-week period
Stepped care	3. Care coordinators reviewed 50% caseload with psychiatrists	At least one psychiatric consultation of the case in each 4-week period
	4. Registry documents current psychiatric medication for 75% of caseload	N/A ^a

^aMedication data were not available for research.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Baseline characteristics of study samples (%)

	Entire sample (n=1,806)	No exposure to VBP (n=556)	At least one month of exposure to VBP (n=1,250)
Demographics			
Female (vs. male)	50.0	46.4	51.6
Age			
18–29	17.8	15.7	18.7
30–39	22.4	18.5	24.1
40–49	31.7	35.3	30.2
50–59	24.5	27.3	23.3
60+	3.6	3.2	3.8
Baseline clinical indicators			
PHQ-9			
10–14	26.0	25.4	26.2
15–19	34.4	31.7	35.6
20+	39.7	43.0	38.2
Comorbid behavioral health conditions			
Anxiety	58.0	53.2	60.2
Bipolar	20.5	18.5	21.4
Cognitive disorder	2.2	1.3	2.6
Psychotic disorder	3.3	3.2	3.3
PTSD	17.4	16.6	17.8
Substance abuse	22.2	22.1	22.2
Suicidal ideation	18.8	19.8	18.4
MHIP eligibility categories^a			
Disability Lifeline	84.8	96.2	79.7
Low-income mothers	8.5	1.4	11.6
Low-income older adults	7.5	4.7	8.8
Uninsured	3.2	1.6	3.8
Veterans	2.4	0.4	3.3
Family members of veterans	0.8	0.5	1.0

^aCategories not mutually exclusive.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Estimated effects of VBP on fidelity to process-of-care elements of collaborative care: adjusted results

Table 3

	Follow-up contacts		Psychiatric consultation		PHQ assessments	
	Entire sample (n=1,781)	DL enrollees (n=1,515)	Entire sample (n=1,781)	DL enrollees (n=1,515)	Entire sample (n=1,781)	DL enrollees (n=1,515)
No VBP	0.57 (0.54, 0.60)	0.59 (0.56, 0.62)	0.14 (0.12, 0.16)	0.15 (0.12, 0.17)	0.48 (0.45, 0.51)	0.51 (0.48, 0.54)
VBP	0.62 (0.60, 0.64)	0.63 (0.60, 0.65)	0.18 (0.17, 0.20)	0.20 (0.18, 0.22)	0.55 (0.53, 0.57)	0.56 (0.54, 0.59)
Marginal effect of VBP	0.05 (0.00, 0.10)	0.04 (0.00, 0.09)	0.04 (0.00, 0.07)	0.05 (0.02, 0.09)	0.07 (0.02, 0.11)	0.05 (0.00, 0.11)

DL, Disability Lifeline; PHQ, Patient Health Questionnaire.

Results are based on mixed-effects linear models with a random intercept at the patient level; sample sizes presented are number of patients included in analysis; 95% confidence interval in parentheses.