Effects of Accountable Care and Payment Reform on Substance Use Disorder Treatment: Evidence from the Initial Three Years of the Alternative Quality Contract

Elizabeth A. Stuart, Ph.D.,
Johns Hopkins Bloomberg School of Public Health

Colleen L. Barry, Ph.D., M.P.P.,
Johns Hopkins Bloomberg School of Public Health

Julie M. Donohue, Ph.D.,
University of Pittsburgh School of Public Health

Shelly F. Greenfield, M.D., M.P.H.,
McLean Hospital, Harvard Medical School

Kenneth Duckworth, M.D.,
Blue Cross Blue Shield of Massachusetts

Zirui Song, M.D., Ph.D.,
Harvard Medical School

Elena M. Kouri, Ph.D.,
Harvard Medical School

Cyrus Ebnesajjad, M.A.,
Johns Hopkins Bloomberg School of Public Health

Robert Mechanic, M.B.A.,
Brandeis University

Location of work and address for correspondence and reprint requests: Johns Hopkins Bloomberg School of Public Health, 624 N Broadway, Room 804, Baltimore, MD 21205; estuart@jhu.edu.

Michael E. Chernew, Ph.D., and
Harvard Medical School

Haiden A. Huskamp, Ph.D.
Harvard Medical School

Abstract

Background and Aims—Global payment and accountable care reform efforts in the US may connect more individuals with substance use disorders (SUD) to treatment. We tested whether such changes instituted under an 'Alternative Quality Contract' (AQC) model within the Blue Cross Blue Shield of Massachusetts' (BCBSMA) insurer increased care for individuals with SUD.

Design—Difference-in-differences design comparing enrollees in AQC organizations with a comparison group of enrollees in organizations not participating in the AQC.

Setting—Massachusetts, USA.

Participants—BCBSMA enrollees aged 13–64 from 2006–2011 (three years prior to and after implementation) representing 1,333,534 enrollees and 42,801 SUD service users.

Measurements—Outcomes were SUD service use and spending and SUD performance metrics. Primary exposures were enrollment in an AQC provider organization and whether the AQC organization did or did not face risk for behavioral health costs.

Findings—Enrollees in AQC organizations facing behavioral health risk experienced no change in the probability of using SUD services (1.64% vs. 1.66%; p=0.63), SUD spending ($2,807 vs. $2,700; p=0.34) or total spending ($12,631 vs. $12,849; p=0.53), or SUD performance metrics (identification: 1.73% vs. 1.76%, p=0.57; initiation: 27.86% vs. 27.02%, p=0.50; engagement: 11.19% vs. 10.97%, p=0.79). Enrollees in AQC organizations not at risk for behavioral health spending experienced a small increase in the probability of using SUD services (1.83% vs. 1.66%; p=0.003) and the identification performance metric (1.92% vs. 1.76%; p=0.007), and a reduction in SUD medication use (11.84% vs. 14.03%; p=0.03) and the initiation performance metric (23.76% vs. 27.02%; p=0.005).

Conclusions—A global payment and accountable care model introduced in Massachusetts USA (in which a health insurer provided care providers with fixed prepayments to cover most or all of their patients' care during a specified time period, incentivizing providers to keep their patients' healthy and reduce costs) did not lead to sizable changes in substance use disorder service use during the first three years following its implementation.

INTRODUCTION

While the global costs of substance use disorders (SUD) are high (1,2) treatment rates worldwide are low. Only one in six problem drug users worldwide, about 4.4 million people, are estimated to receive required drug treatment, and treatment rates vary substantially by region (3). In the US, fewer than 20% of individuals with drug use disorders and 10% with alcohol use disorders receive treatment (4). In the U.S., SUD is often undetected and goes untreated in primary care settings, and likewise, medical care needs of individuals with SUD are not properly addressed in the specialty SUD treatment sector (5). This is due in large part

to the ongoing separation of the delivery and financing of SUD treatment from the rest of the medical system in the US. It is also due to the fee-for-service payment system common in the U.S., where a provider is reimbursed for each service delivered with no financial incentives to encourage coordinating care across multiple providers or choosing to provide only the highest value services within a fixed budget (6).

Private insurers and government insurance programs in the U.S. are beginning to experiment with new financing and delivery system models to encourage the provision of high value services and emphasize care coordination and improved care integration. Instead of fee-for-service payment, these payment delivery system reform models typically hold large provider organizations accountable for spending on all health care services used by a defined population over a longer (e.g., 5 or 10 years) time horizon. So-called global payment, for example, involves a provider organization being allocated a fixed budget for all care delivered to its patients by a health insurance company. These models also generally place greater accountability on provider organizations for the quality of services provided through the use of financial incentives for meeting specified performance metrics (7).

This paper investigates the effects of the first model of global payment and accountable care implemented in the U.S., the Alternative Quality Contract (AQC). The AQC was adopted by Blue Cross Blue Shield of Massachusetts (BCBSMA) in 2009. The AQC pays large provider organizations a risk-adjusted global payment for all primary and specialty care provided to a defined population of patients over a five-year period, with bonuses in the initial years of up to 10% of the global budget based on performance on a set of 64 quality metrics (Appendix A). Research has shown that AQC implementation was associated with lower overall health care spending growth (primarily driven by a shift to lower-cost providers and a reduction in use of diagnostic and testing services) and improved performance on most performance metrics (8,9). Importantly, none of the performance metrics in the initial AQC five-year contracting period or in more recent contracts have been specific to SUD treatment. In addition, financial risk for behavioral health service use was included (i.e., payment for behavioral health services was included in the global payment) in only a subset of contracts with AQC organizations (5 of 12) established in the initial three years of the AQC (2009 to 2011).

The effects of accountable care and global payment models like the AQC on care for individuals with SUD are unknown. Better integration of SUD and general medical service delivery and financing under the AQC model of global payment and accountable care could improve SUD treatment rates and quality of care (10–12). Since AQC organizations face a global budget for all services received by the patients they serve, AQC organizations have an incentive to improve the integration and efficiency of care to try to prevent high health care expenditures by their patients, including those with chronic SUD and co-occurring medical and mental health conditions. The AQC's five-year contracts (in contrast with the more typical 12 month contracts between provider organizations and health plans in the U.S.) provide incentives to identify and treat SUD using high value, cost-effective interventions. Under the AQC, primary care physicians (PCPs) are given greater flexibility to coordinate the care received by their patients, with an eye toward reducing the burden of disease over the longer term given the five-year contracts. On the other hand, organizations at risk for a

population's total health care spending could avoid enrolling or restrict services for individuals seeking treatment for SUDs since this group tends to have higher total health care expenditures than other individuals (13).

Both public (the Medicaid and Medicare programs, government programs that provides insurance to lower-income individuals and elderly and disabled individuals, respectively) and other private insurance plans in the U.S. are currently considering or have recently implemented models that are similar to the AQC. It is critical to assess the early experience of the AQC in order to understand the potential impact of these new models on care for individuals with SUD. This study tested hypotheses regarding whether the AQC affected: 1) use of SUD services among enrollees, 2) health care spending (overall and SUD-specific), and 3) three SUD-specific performance metrics for enrollees in AQC organizations (14), as compared to enrollees in organizations that did not participate in the AQC, during the first 3 years of the AQC model's implementation (2009–2011). We also tested whether the effects varied across enrollees in AQC organizations that faced risk for behavioral health spending and enrollees in AQC organizations that did not face behavioral health risk.

METHODS

Data

This study uses six years of BCBSMA inpatient, outpatient and pharmacy claims data (2006–2011). The population includes all individuals aged 13–64 years who were continuously enrolled in a BCBSMA health maintenance organization (HMO) or point of service (POS) plan for at least one calendar year during the study period. HMO and POS health plan types involve care that is typically managed more stringently by the commercial insurance company than other types of health plan products. Within the HMO and POS framework, primary care for the enrollees was provided in Massachusetts but the enrollees could have lived outside of Massachusetts. All BCBSMA HMO and POS plan enrollees are required to select a PCP upon enrollment. Participation in the AQC was determined by whether an enrollee's PCP belongs to an organization that has entered an AQC contract; 7 organizations entered in 2009, 4 in 2010, and 1 in 2011. The comparison group for analyses thus includes both individuals in organizations that never entered the AQC and enrollees in organizations that later adopted the AQC. The unit of observation was the person-year. We included only those person-years with 12 months of enrollment in medical, behavioral, and pharmacy benefits managed by BCBSMA.

Study Design

The study uses a comparative interrupted time series design, also known as a difference-in-differences (DD) design, to estimate the effects of the AQC, which tests whether the trends in the outcomes of interest change before and after organizations enter the AQC, and in particular whether the temporal changes in trends differ between organizations that did and did not actually enter the AQC. This design is commonly used in policy evaluations (8, 9, 15–17); and is valid under an assumption that the differences in trends between AQC and non-AQC organizations in the pre-AQC time period would have continued in the same way

had the AQC never been initiated. Propensity score weights were used to ensure that the AQC and non-AQC groups were similar to each other on observed covariates in each year.

Identifying SUD Service Use

We employed a common approach to identifying SUD service users in insurance claims (18, 19). SUD diagnoses were identified using the *International Classification of Diseases, 9th Revision, Clinical Modification* (ICD-9-CM) codes, in particular 291–292, 303–305.0, and 305.2–305.9. Inpatient care was coded as SUD-related if the last primary diagnosis and the majority of all primary diagnoses in the inpatient admission were SUD diagnoses (as in (18)). Outpatient visits were classified as SUD if there was a SUD primary or secondary diagnosis on the claim (as in (20)); secondary diagnoses were included due to concerns that SUD has been systematically undercoded in outpatient claims. A person-year was considered a SUD service user if he or she had at least one SUD-related inpatient stay or outpatient visit during the calendar year. We considered a person-year with either of the following to have used a SUD medication: 1) pharmacy claim for naltrexone, disulfiram, acamprosate, methadone, or buprenorphine; or 2) outpatient claim with a HCPCS/CPT code for administration of methadone or buprenorphine.

Outcome Measures

We estimated the effect of the AQC on: (1) the probability of SUD service use; (2) average SUD spending and average total health care spending among SUD service users; (3) the probability and quantity of specific types of SUD services used; and (4) HEDIS-based performance measures. Among SUD service users, we examined the probability of using detoxification services, inpatient SUD services, and SUD-specific medications. We did not model the probability of using outpatient SUD services because nearly all SUD service users received outpatient services. We examined the number of outpatient SUD visits among outpatient SUD users and the number of inpatient SUD days among inpatient SUD users. In the primary analyses detoxification services are included in the inpatient and outpatient counts, and in spending; a sensitivity analysis (results in the Appendix) examined the same outcomes but with claims for detoxification services excluded.

We also examined three Healthcare Effectiveness Data and Information Set (HEDIS)-based SUD performance measures (see Appendix B for details). HEDIS is a tool used by more than 90 percent of health plans in the U.S. to measure performance on important dimensions of care and service. In this case, we used three HEDIS measures – SUD identification, treatment initiation, and treatment engagement – originally developed by the Washington Circle to capture how well health plans were performing with regard to SUD treatment (21). We defined *identification* as the share of all enrollees who had a new SUD claim within a calendar year. We defined *treatment initiation* as the share of enrollees with a new episode of SUD treatment who initiated treatment within 14 days of their initial diagnosis. We defined *treatment engagement* as the share of enrollees with a new episode of SUD treatment who received at least two SUD services within 30 days of initial diagnosis.

Statistical Analyses

For each outcome we examined two primary models within the difference-in-differences design. The first estimates the overall effect of the AQC by fitting a model of the outcome as a function of AQC status (which is the primary predictor of interest), and a set of control variables. The second model estimates effects separately among enrollees in organizations that faced risk for behavioral health and those that did not (both relative to enrollees in organizations that did not participate in the AQC), by replacing the AQC indicator by a three-level variable to identify organizations as non-AQC, AQC with behavioral health risk, or AQC without behavioral health risk. In supplementary analyses we also examined whether the effects varied by how long the organizations had participated in the AQC by including entry year by AQC interaction terms. All models controlled for gender; age categories (13–17, 18–27, 28–37, 38–47, 48–57, 58–64); the interaction of age and gender; risk score; whether the enrollee's provider organization entered the AQC in 2009, 2010 or 2011; and calendar year. The risk score was calculated by BCBSMA from current-year diagnoses (including comorbid mental health conditions), claims and demographic information and is based on the diagnostic-cost-group (DxCG) scoring system (Verisk Health) (22). Given that the outcomes of patients being seen in the same physician practices may be correlated, the standard errors accounted for clustering at the practice level by using Taylor series variance estimation, implemented using the svy commands in Stata 13.1 (23). The resulting estimates can be interpreted as the average annual outcome if all enrollees were subject to the AQC versus if all enrollees were not subject to the AQC during the study period.

To adjust for potential differences in enrollee characteristics across the intervention and comparison groups, we used propensity score weighting (24). These "inverse probability of treatment weights" weight each of the groups (comparison enrollees, enrollees in AQC organizations with behavioral health risk, enrollees in AQC organizations without behavioral health risk) up to the combined sample of all three groups, thus making each of them comparable to each other. The propensity score model included age group, gender, and the risk score. Note that these variables were also included as predictors in the outcome regression models; a large body of research has shown the benefits of including covariates in both the propensity score and outcome models (25–28). The weights are constructed separately for each calendar year (2009–2011); in the pre-AQC period (2006–2008) all weights are 1. It could be problematic to adjust for variables that are affected by the AQC (in the post-AQC years of 2009–2011); however, it is unlikely that the case mix of enrollees served by providers with respect to these characteristics would change as a result of the AQC, and it is important to adjust for them to remove confounding of the effects from any differences in case-mix across the groups.

Two-part models were used to examine changes in the probability of any SUD service use and the amount of spending, conditional on SUD treatment, attributable to the AQC. The model's first stage used a logistic model while the second stage used a linear model (as in (29)). The probability of receiving specific types of services and the share of enrollees meeting HEDIS-based performance metrics were modeled using logistic regression. The quantity of specific types of services used was modeled using negative binomial models.

Model diagnostics were used to determine the primary analysis models; however, we also conducted sensitivity analyses to check the robustness of our findings, including modeling spending using a log-gamma model and a log-transformed model. For the quantity measures, we conducted sensitivity analyses using Poisson models instead of negative binomial models. Finally, we estimated models without propensity score weights. All analyses were run in Stata 13.1 (30).

RESULTS

Descriptive statistics of SUD treatment users at the person-year level are provided in Table 1. SUD service use among BCBSMA enrollees was low, with 1.5% of the comparison group and 1.8% of the AQC group using any SUD services in a calendar year. Nearly two-thirds of SUD users in a given year also received services for a mental health condition in that year. The two most common SUD diagnoses were alcohol and opioid use disorders.

The results from the difference-in-differences models estimating the effects of the AQC on the probability of using SUD services, and SUD and total health care spending among SUD service users are shown in Table 2 and Figure 1. Estimates indicate that the AQC had no overall effect on the probability of using SUD services, average spending on SUD services, or average total health care spending among SUD service users. We detected a slight increase in the probability of SUD service use among enrollees in AQC organizations that did not face risk for behavioral health relative to the comparison group.

Table 3 examines the effect of the AQC on the probability of using specific types of SUD services and, among those using services, the quantity of specific services used. The AQC had no effect on the probability of receiving detoxification services, SUD-related inpatient use, or SUD medication use among SUD service users (except a potential decrease in SUD medication use among enrollees in organizations that did not face behavioral health risk a marginally significant increase in the number of inpatient days). The AQC had no effect on the amount of detox received, the number of outpatient SUD services among users, or on inpatient use.

Table 4 indicates no overall effects of the AQC on HEDIS-based SUD performance measures. Enrollees in AQC organizations that did not face behavioral health risk had slightly higher rates of identification (consistent with results on SUD service use in Table 2), but lower rates of treatment initiation than enrollees in non-AQC organizations attributable to being in the AQC. We found no effect of the AQC on treatment engagement.

Results from the sensitivity analyses (Appendix C) were qualitatively similar across all analyses, including analyses that excluded all claims for detoxification and unweighted analyses, with the exception that unweighted analyses indicated a slight increase in the number of inpatient days due to the AQC (9.89 vs. 8.49 days; p=0.01), both overall and for both the risk group (9.99 vs. 8.49 days; p=0.01) and the no-risk group (9.80 vs. 8.49 days; p=0.03), and a decrease in the number of outpatient visits among enrollees in AQC groups that did not face behavioral health risk (7.94 vs. 8.44; p=0.01). There was generally no evidence of effects varying by year of implementation.

DISCUSSION

Despite growing interest in global payment and accountable care models among both public and private payers in the U.S. and their potential for improving care for SUDs, we found almost no evidence that the AQC affected SUD performance metrics or spending on enrollees with SUD treatment in the first three years of the contract. We documented a slight increase in the probability of receiving any SUD services among enrollees in AQC organizations that did not face risk for behavioral health care use (and a parallel effect in the HEDIS-based SUD treatment identification measure). While previous research has documented slower growth in total health care spending among the full population of AQC enrollees (8, 9, 29), there was no change in total health care spending or SUD-treatment specific spending among SUD service users. Those overall savings for the full population were driven in large part by referrals to lower-cost providers in the early years and reductions in volume of procedures, imaging, and tests in later years (9, 29), which may be harder to achieve in the SUD context given the shortage of substance use treatment providers and general challenges to care coordination and adherence for patients who suffer from SUDs. It is notable that fewer than 2% of this sample used any SUD services in a calendar year. Given the longstanding under-diagnosis and under-treatment of patients with SUDs (20), a reduction in the growth of SUD spending may be a less important aim for the provider organizations compared with the goal of improving access to evidence-based treatments.

Limited focus on improving integration and quality of SUD treatment during the early years of the AQC is consistent with results from recent surveys of accountable care organizations (31) and of substance abuse treatment (SAT) organizations (32) in the United States. Only 14% of accountable care organizations surveyed reported that integration of behavioral health and primary care was nearly or fully complete (31), while only 15% of SAT organizations had signed agreements to be part of an ACO (32). It is also consistent with knowledge about the implementation of the AQC program itself; AQC organization leaders reported little focus on this area during the AQC's first three years as they struggled to reorient their systems in response to the AQC. Likewise, leaders of specialty behavioral health provider organizations in Massachusetts have indicated a low level of awareness of the AQC, and no substantial changes in referral patterns or contracting practices due to the AQC (33).

In recent years (after our study period) several AQC organizations have initiated efforts to improve the integration of behavioral health and other services, including using behavioral health case managers or social workers on clinical teams that work to integrate care for behavioral health and other chronic conditions for complex patients (33). In addition, in 2012, BCBSMA took a leadership role nationally in addressing the opioid epidemic plaguing the U.S. by changing its policies governing coverage for prescription opioid medications. Future work should investigate the effects of these more recent initiatives on SUD treatment, to examine if more focused attention and efforts led to effects on such care (34–35).

There are likely still barriers, however, to efforts to improve SUD coordination and quality. First, the U.S. 42 Code of Federal Regulations (or 42 CFR) restricting the disclosure and use of patient SUD records represents a major barrier to SUD integration efforts, making it difficult for health care providers to share information (such as on current or past alcohol or drug use problems or SUD service use) with providers or coordinate the services received by patients.

Second, the lack of SUD performance measures in the AQC contracts, and in the field more generally, may have also served as a barrier to efforts to improve SUD quality and coordination, as AQC organizations likely focused on those areas of care that are measured and incented under the contract. Global payment models rely on financial incentives for meeting performance measures to protect against under-provision or stinting on care. While the lack of meaningful, validated quality measures for mental health treatment generally has been noted as a problem for improving quality of mental health services (36), the lack of good SUD measures is perhaps even more striking. As risk-based payment models expand, a broader set of SUD measures is needed to ensure high quality SUD care. However, while such measures may help to enhance integration and quality, provider organizations may be challenged in implementing the measures while balancing privacy needs of patients and the federal privacy protections of the 42 CFR.

Finally, many locations in the U.S., including Massachusetts, have a shortage of substance abuse treatment providers with strong linkages with large provider groups; this too limits the ability of models such as the AQC to improve care quality and coordination.

Several limitations should be noted. First, this analysis is focused on a global payment and accountable care model implemented by a single private insurer in the U.S.; thus, results may not be generalizable to other reforms implemented by either public or private payers. Second, while AQC organizations with contracts that included behavioral health risk and those that did not include it did not appear to systematically differ in obvious ways (e.g., size, level of integration), it is possible that there are unobservable differences. However, we did not observe differences in pre-intervention trends in outcomes between each AQC cohort and the comparison group. In addition, the analysis adjusts for any differences in trends during the pre-AQC period, using an assumption that they would have persisted over time. Third, we focus our analyses on the first three years of AQC implementation, so we are unable to examine the long-term impacts of this change. As of 2011, approximately 44% of BCBSMA HMO and POS enrollees were part of the AQC; we viewed this as a tipping point in terms of having a viable BCBSMA comparison group. However, this means we are not able to determine whether adoption of the AQC led to changes affecting individuals with SUD in the years following the end of our study period. It is not uncommon for changes in outcomes to occur years after policy adoption given inherent challenges associated with altering well-established provider practice patterns. Finally, claims data, which provide comprehensive information on service use and spending, are limited in their ability to capture clinical detail that may be relevant for understanding quality of care. Nevertheless, this study provides the first information about the impact on SUD treatment of risk-based payment models that are increasingly being considered and implemented by both public and

private payers in the U.S. as an effort to improve the efficiency and quality of service delivery.

Despite concerns about under-provision of SUD services under global payment, this study found that SUD service use and spending did not decline as a result of the AQC. However, it is also clear that the AQC did not spur major improvements in care for SUD during the contract's initial years. As AQC organizations begin to put in place initiatives to improve behavioral health integration, it will be critical to continue to track whether these changes translate into meaningful change in service use, health care spending, or care quality for persons with SUD treatment needs.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

REFERENCES

1. Harwood, HJ. International Symposium on the Economic Costs of Substance Abuse. Banff (Canada): 1994 May 11. Analytical principles and issues in making cost of illness estimates for substance abuse; p. 11-13.

2. World Health Organization. Economic and social costs of alcohol use Global Status Report on Alcohol 2004. Geneva: 2004. p. 65-66.

3. International Narcotics Control Board. Report of the International Narcotics Control Board for 2014. United Nations Publications; 2015.

4. Substance Abuse and Mental Health Services Administration. Results from the 2012 National Survey on Drug Use and Health: Summary of National Findings, NSDUH Series H-46, HHS Publication No. (SMA) 13-4795. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2013.

5. Bouchery E, Harwood HJ, Dilonardo J, Vandivort-Warren R. Type of health insurance and the substance abuse treatment gap. J Subst Abuse Treat. 2012; 42(3):289–300. [PubMed: 22119184]

6. Barry CL, Huskamp HA. Moving beyond parity—mental health and addiction care under the ACA. N Engl J Med. 2011; 365(11):973–975. [PubMed: 21848453]

7. Gosline, A.; Rodman, E. Summary of Chapter 224 of the Acts of 2012. Boston: Blue Cross Blue Shield of Massachusetts Foundation; 2012 Sep. http://bluecrossmafoundation.org/sites/default/files/download/publication/Chapter%20224%20summary_1.pdf. (Archived by WebCite® at http://www.webcitation.org/6j1jFiL6m) [Accessed 15 July 2016]

8. Song Z, Safran DG, Landon BE, Landrum MB, He Y, Mechanic RE, et al. The 'Alternative Quality Contract,' based on a global budget, lowered medical spending and improved quality. Health Aff (Millwood). 2012; 31(8):1885–1894. [PubMed: 22786651]

9. Song Z, Rose S, Safran DG, Landon BE, Day MP, Chernew ME. Changes in health care spending and quality 4 years into global payment. N Engl J Med. 2014; 371(18):1704–1714. [PubMed: 25354104]

10. Crowley R, Kirschner N. Health and Public Policy Committee of the American College of Physicians. The Integration of Care for Mental Health, Substance Abuse, and Other Behavioral

Health Conditions into Primary Care: An American College of Physicians Position Paper. Ann Intern Med. 2015; 163(4):298–299. [PubMed: 26121401]

11. Andrews C, Abraham A, Grogan CM, Pollack HA, Bersamira C, Humphreys K, Friedmann P. Despite resources from the ACA, most states do little to help addiction treatment programs implement health care reform. Health Aff (Millwood). 2015; 34(5):828–835. [PubMed: 25941285]

12. Buck JA. The looming expansion and transformation of public substance abuse treatment under the Affordable Care Act. Health Aff (Millwood). 2011; 30(8):1402–1410. [PubMed: 21821557]

13. Frank, RG.; McGuire, TG. Economics and Mental Health. In: Culyer, AJ.; Newhouse, JP., editors. Handbook of Health Economics. Oxford: Elsevier; 2000. p. 893-954.

14. McCorry F, Garnick DW, Bartlett J, Cotter F, Chalk M. Developing performance measures for alcohol and other drug services in managed care plans. Washington Circle Group. Jt Comm J Qual Improv. 2000; 26(11):633–643.

15. Biglan A, Ary D, Wagenaar AC. The value of interrupted time-series experiments for community intervention research. Prev Sci. 2000; 1(1):31–49. [PubMed: 11507793]

16. Mason T, Sutton M, Whittaker W, McSweeney T, Millar T, Donmall M, et al. The impact of paying treatment providers for outcomes: difference-in-difference analysis of the 'payment by results for drugs recovery' pilot. Addiction. 2015; 110(7):1120–1128. [PubMed: 26058447]

17. Nakamura R, Suhrcke M, Pechey R, Morciano M, Roland M, Marteau TM. Impact on alcohol purchasing of a ban on multi-buy promotions: a quasi-experimental evaluation comparing Scotland with England and Wales. Addiction. 2014; 109(4):558–567. [PubMed: 24251415]

18. Goldman HH, Frank RG, Burnam MA, Huskamp HA, Ridgely MS, Normand SL, et al. Behavioral health insurance parity for federal employees. N Engl J Med. 2006; 354(13):1378–1386. [PubMed: 16571881]

19. McConnell KJ, Gast SH, Ridgely MS, Wallace N, Jacuzzi N, Rieckmann T, et al. Behavioral health insurance parity: does Oregon's experience presage the national experience with the Mental Health Parity and Addiction Equity Act? Am J Psychiatry. 2012; 169(1):31–38. [PubMed: 21890792]

20. Greenfield SF, Azzone V, Huskamp H, Cuffel B, Croghan T, Goldman W, et al. Treatment for substance use disorders in a privately insured population under managed care: Costs and services use. J Subst Abuse Treat. 2004; 27(4):265–275. [PubMed: 15610828]

21. The Washington Circle. [Accessed 30 July 2016] Specification of performance measures for the identification, initiation, and engagement of alcohol and other drug services. 2004. http://washingtoncircle.org/pdfs/FinalSpecs202004_03-30.pdf (Archived by WebCite® at http://www.webcitation.org/6jP0HJngU)

22. Pope GC, Kautter J, Ellis RP, Ash AS, Ayanian JZ, Lezzoni LI, et al. Risk adjustment of Medicare capitation payments using the CMS-HCC model. Health Care Financ Rev. 2004; 25(4):119–141. [PubMed: 15493448]

23. StataCorp. [Accessed 15 July 2016] Stata Survey Data Reference Manual: Release 13. 2013. Available at: http://www.stata.com/manuals13/svy.pdf (Archived by WebCite® at http://www.webcitation.org/6j1jQFx1Z)

24. Rosenbaum PR, Rubin DB. The central role of the propensity score in observation studies for causal effects. Biometrika. 1983; 70:41–55.

25. Stuart EA. Matching methods for causal inference: A review and a look forward. Stat Sci. 2010; 25(1):1–21. [PubMed: 20871802]

26. Bang H, Robins JM. Doubly robust estimation in missing data and causal inference models. Biometrics. 2005; 61(4):962–973. [PubMed: 16401269]

27. Ho DE, Imai K, King G, Stuart EA. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. Political Analysis. 2007; 15:199–236.

28. Rubin DB. The use of matched sampling and regression adjustment to remove bias in observational studies. Biometrics. 1973; 29(1):185–203.

29. 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. N Engl J Med. 2011; 365(10):909–918. [PubMed: 21751900]

30. StataCorp.. Stata Statistical Software: Release 13. College Station, TX: StataCorp LP; 2013.

31. Lewis VA, Colla CH, Tierney K, Van Citters AD, Fisher ES, Meara E. Few ACOs pursue innovative models that integrate care for mental illness and substance abuse with primary care. Health Aff (Millwood). 2014; 33(10):1808–1816. [PubMed: 25288427]

32. D'Aunno T, Friedmann PD, Chen Q, Wilson DM. Integration of Substance Abuse Treatment Organizations into Accountable Care Organizations: Results from a National Survey. J Health Polit Policy Law. 2015; 40(4):797–819. [PubMed: 26124307]

33. Barry CL, Stuart EA, Donohue J, Greenfield S, Kouri E, Duckworth K, Song C, Mechanic R, Chernew M, Huskamp HA. Effects of Payment Reform and Accountable Care on Mental Health: Early Experience with the "Alternative Quality Contract" in Massachusetts. Health Aff (Millwood). 2015; 34(12):2077–2085. [PubMed: 26643628]

34. Marbury, D. [Accessed 30 July 2016] A health plan tackles opioid addiction: Program reduces risk of misuse. Managed Healthcare Executive. 2015 Mar 12. http://managedhealthcareexecutive.modernmedicine.com/managed-healthcare-executive/news/health-plan-tackles-opioid-addiction?page=full (Archived by WebCite® at http://www.webcitation.org/6jOziwZIJ)

35. Weisner CM, Chi FW, Lu Y, Ross TB, Wood SB, Hinman A, et al. Examination of the effects of an intervention aiming to link patients receiving addiction treatment with health care: the LINKAGE clinical trial. JAMA Psychiatry. [Epub ahead of print 2016 Jun 22].

36. Goldman ML, Spaeth-Rublee B, Pincus HA. Quality Indicators for Physical and Behavioral Health Care Integration. JAMA. 2015; 314(8):769–770. [PubMed: 26043185]
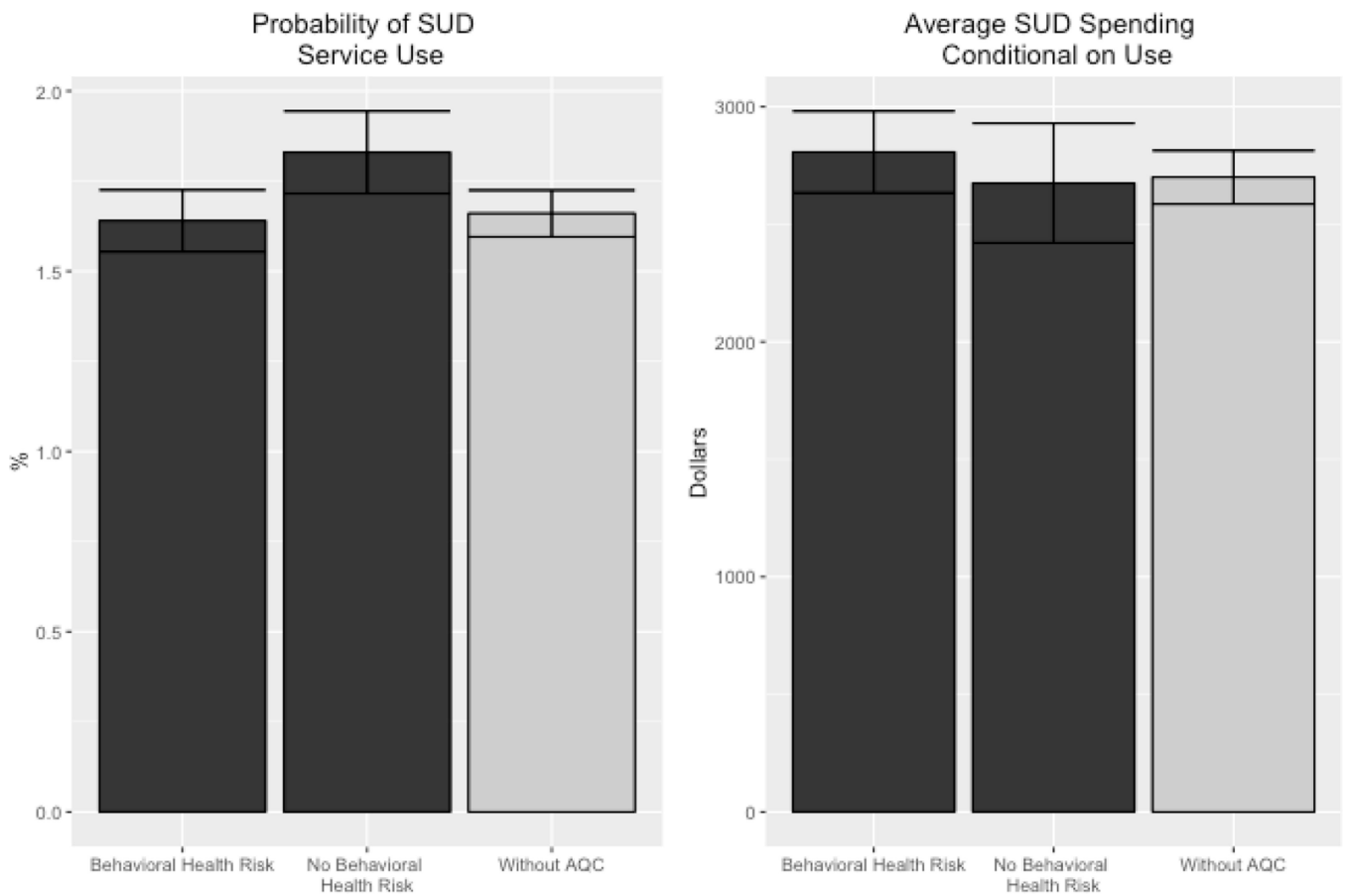
Figure 1.
Effects of the Alternative Quality Contract (AQC) on the Probability of SUD Service Use
and on SUD Spending among SUD Service Users.
Note: Results from Table 2. Models fit using two-part models with linear models for
spending. Models were propensity-score weighted and adjusted for gender, age category,
risk score, year and AQC cohort. Dollar values adjusted for inflation. Difference-in-
differences estimation used to account for secular trends.

Table 1

Unadjusted Descriptive Statistics for Enrollees Using Substance Use Disorder Services in the Comparison and AQC Groups (With and Without Behavioral Health Risk) Across Person-Years, 2006–2011

	Comparison Group		AQC Group		AQC Group with Behavioral Health Risk		AQC Group without Behavioral Health Risk	
N (person-years)	50,576		10,817		4,608		6,209	
Female (%, n)	35.5%	17,929	36.0%	3,895	34.1%	1,569	37.5%	2,326
Age (%, n)								
13–17	4.91%	2,483	3.2%	342	2.8%	130	3.4%	212
18–27	23.8%	12,018	25.4%	2,748	25.5%	1,174	25.4%	1,574
28–37	14.8%	7,489	15.5%	1,678	15.7%	721	15.4%	957
38–47	23.0%	11,606	20.7%	2,241	21.1%	971	20.5%	1,270
48–57	23.9%	12,061	24.6%	2,663	25.4%	1,171	24.0%	1,492
58–64	9.7%	4,919	10.6%	1,145	9.6%	441	11.3%	704
Risk score (mean)	3.07		3.10		3.16		3.06	
Mental health disorder (%, n)	67.6%	34,188	68.7%	7,427	66.8%	3,077	70.1%	4,350
SUD diagnosis groups								
Alcohol use disorders	65.8%	33,270	65.4%	7,076	63.5%	2,924	66.9%	4,152
Opioid use disorders (%, n)	20.6%	10,423	23.9%	2,582	27.3%	1,256	21.4%	1,326
Stimulant use disorders (%, n)	0.8%	424	1.0%	106	1.0%	46	1.0%	60
Cocaine use disorders	6.8%	3,460	5.2%	557	5.4%	250	5.0%	307
Other substance use disorder (%, n)[1]	37.4%	18,928	37.2%	4,028	38.8%	1,786	36.1%	2,242
SUD service use in a calendar year								
Any outpatient SUD visits, among SUD service users (%, n)	99.6%	50,396	99.7%	10,782	99.6%	4,591	99.7%	6,191

	Comparison Group		AQC Group		AQC Group with Behavioral Health Risk		AQC Group without Behavioral Health Risk	
Any detoxification services, among SUD service users (%, n)	9.1%	4,598	9.8%	1,063	10.9%	503	9.0%	560
Any SUD medication use, among SUD service users (%, n)	13.2%	6,670	14.1%	1,520	16.2%	746	12.5%	774
Any inpatient SUD days, among SUD service users (%, n)	4.44%	2,245	5.17%	559	5.43%	250	4.98%	309
Average # of outpatient SUD visits, among outpatient SUD	8.48	50,396	7.94	10,782	8.52	4,591	7.51	6,191
Average # of inpatient SUD days, among inpatient SUD service users	8.49	2,245	9.83	559	10.06	250	9.64	309
Average # of inpatient SUD stays, among inpatient SUD service users	1.28	2,245	1.30	559	1.28	250	1.31	309
Spending in a calendar year								
Average total SUD spending among SUD service users (mean $, n)	$2,526	50,576	$2,874	10,817	$3,034	4,608	$2,755	6,209
Average total health care spending among SUD service users (mean $, n)	$12,546	50,576	$12,821	10,817	$12,899	4,608	$12,763	6,209

[1] Other SUD use includes use of cannabis, hallucinogens, sedatives, hypnotics, anxiolytics, antidepressants, mixed, and unspecified drug use, as well as drug induced mental disorders.

Table 2

Adjusted Annual Changes in Probability of Substance Use Disorder (SUD) Service Use and Average SUD and Total Health Care Spending Conditional on Use in the Alternative Quality Contract (AQC), 2006–2011 (N=3,779,879 person years; N=61,393 SUD service-user person-years)[1]

	With AQC	Without AQC	Difference	P-value	[95% Conf. Interval]
Probability of Use (%)					
Overall	1.72%	1.67%	0.04%	0.27	−0.03% to 0.11%
AQC Behavioral Health Risk	1.64%	1.66%	−0.03%	0.63	−0.13% to 0.08%
AQC Behavioral Health No Risk	1.83%	1.66%	0.16%	0.003 *	0.06% to 0.26%
Average Total SUD Spending (conditional on being an SUD service user) ($)					
Overall	$2,747	$2,689	$58	0.54	−$124 to $239
AQC Behavioral Health Risk	$2,807	$2,700	$108	0.34	−$112 to $328
AQC Behavioral Health No Risk	$2,675	$2,700	−$25	0.85	−$285 to $235
Average Total Health Care Spending (conditional on being an SUD service user) ($)					
Overall	$12,762	$12,872	−$110	0.70	−$678 to $458
AQC Behavioral Health Risk	$12,631	$12,849	−$218	0.53	−$892 to $455
AQC Behavioral Health No Risk	$12,917	$12,849	$68	0.83	−$560 to $695

[1]Models fit using two-part models with linear models for spending. Models were propensity-score weighted and adjusted for gender, age category, risk score, year and AQC cohort. Dollar values adjusted for inflation. Difference-in-differences estimation used to account for secular trends.

*
p<=0.05

Table 3

Adjusted Annual Changes in the Type and Quantity of SUD Services Used in the Alternative Quality Contract (AQC), 2006–2011 (N=3,779,879 person years; N=61,393 SUD service-user person-years)[1]

	With AQC	Without AQC	Difference	P-value	[95% Conf. Interval]
Probability of any detoxification services (conditional on being an SUD service user)					
Overall	9.55%	9.36%	0.18%	0.74	−0.89% to 1.25%
AQC Behavioral Health Risk	10.40%	9.52%	0.84%	0.28	−0.68% to 2.37%
AQC Behavioral Health No Risk	8.53%	9.52%	−1.03%	0.17	−2.51% to 0.45%
Probability of any SUD medication use (conditional on being an SUD service user)					
Overall	13.86%	13.72%	0.14%	0.85	−1.32% to 1.60%
AQC Behavioral Health Risk	15.56%	14.03%	1.44%	0.13	−0.43% to 3.32%
AQC Behavioral Health No Risk	11.84%	14.03%	−2.30%	0.03 *	−4.43% to −0.18%
Probability of any SUD-related inpatient use (conditional on being an SUD service user)					
Overall	4.95%	4.98%	0.03%	0.94	−0.85% to 0.78%
AQC Behavioral Health Risk	5.22%	5.04%	0.17%	0.72	−0.78% to 1.13%
AQC Behavioral Health No Risk	4.60%	5.04%	−0.46%	0.39	−1.51% to 0.59%
Average # of outpatient SUD service visits (conditional on being an outpatient SUD service user)					
Overall	8.09	8.36	−0.26	0.58	−1.19 to 0.67
AQC Behavioral Health Risk	8.46	8.41	0.04	0.94	−1.05 to 1.14
AQC Behavioral Health No Risk	7.66	8.41	−0.78	0.12	−1.76 to 0.21
Average # of inpatient days (conditional on being an inpatient SUD service user)					
Overall	9.73	8.60	1.24	0.06	−0.04 to 2.32
AQC Behavioral Health Risk	9.92	8.63	1.28	0.06	−0.07 to 2.63
AQC Behavioral Health No Risk	9.47	8.63	0.85	0.25	−0.59 to 2.29

[1] Probability of use models fit using logistic regression. Visit quantity models fit using negative binomial models. All models propensity score weighed and adjusted for gender, age category, risk score, year and AQC cohort. Difference-in-differences estimation used to account for secular trends.

*
p<=0.05

Table 4

HEDIS SUD Quality Measures in the Alternative Quality Contract (AQC), 2006–2011 (N=3,779,879 person years)[1]

	With AQC	Without AQC	Difference	P-value	[95% Conf. Interval]
SUD Identification					
Overall	1.81%	1.77%	0.04%	0.35	−0.04% to 0.11%
AQC Behavioral Health Risk	1.73%	1.76%	−0.03%	0.57	−0.14% to 0.08%
AQC Behavioral Health No Risk	1.92%	1.76%	0.15%	0.007 *	0.04% to 0.26%
SUD Treatment Initiation					
Overall	25.99%	26.70%	−0.71%	0.36	−2.24% to 0.81%
AQC Behavioral Health Risk	27.86%	27.02%	0.82%	0.50	−1.57% to 3.21%
AQC Behavioral Health No Risk	23.76%	27.02%	−3.34%	0.005 *	−5.68% to −1.00%
SUD Treatment Engagement					
Overall	10.42%	10.83%	−0.41%	0.46	−1.50% to 0.68%
AQC Behavioral Health Risk	11.19%	10.97%	0.21%	0.79	−1.38% to 1.80%
AQC Behavioral Health No Risk	9.51%	10.97%	−1.50%	0.08	−3.20% to 0.19%

[1] Logistic regression models propensity-score weighted and adjusted for gender, age category, risk score, year and AQC cohort. Difference-in-differences estimation used to account for secular trends. *Identification* refers to the share of all enrollees who had a new SUD claim within a calendar year. *Treatment initiation* refers to the share of enrollees with a new episode of SUD treatment who initiated treatment within 14 days of their initial diagnosis. *Treatment engagement* refers to the share of enrollees with a new episode of SUD treatment who received at least two SUD services within 30 days of their initial diagnosis. See Appendix B for details.

*
p<=0.05