Published in final edited form as:

J Subst Abuse Treat. 2017 November ; 82: 93–101. doi:10.1016/j.jsat.2017.09.007.

Influencing Quality of Outpatient SUD Care: Implementation of Alerts and Incentives in Washington State

Deborah W. Garnick, ScD¹, Constance M. Horgan, ScD¹, Andrea Acevedo, PhD^{1,3}, Margaret T. Lee, PhD¹, Lee Panas, MS¹, Grant A. Ritter, PhD¹, Kevin Campbell, DrPH², and Jason Bean-Mortinson⁴

¹Institute for Behavioral Health, The Heller School for Social Policy and Management, Brandeis University

²The Division of Behavioral Health and Recovery, Washington State Behavioral Health Administration

³Department of Community Health, Tufts University

⁴Thurston Mason Behavioral Health Organization

Abstract

Financial incentives for quality improvement and feedback on specific clients are two approaches to improving the quality of treatment for individuals with substance use disorders. We examined the impacts of these interventions in Washington State by randomizing outpatient substance use treatment agencies into intervention and control groups. From October 2013 through December 2015, agencies could earn financial incentives for meeting performance goals incorporating both achievement relative to a benchmark and improvement from agencies' own baselines. Weekly feedback was e-mailed to agencies in the alert or alert plus incentives arms. Difference-in-difference regressions controlling for client and agency characteristics showed that none of the interventions significantly affected client engagement after outpatient admissions, overall or for sub-groups based on race/ethnicity, age, rural residence, or agency baseline performance. Treatment agencies offered insights related to several themes: delivery system context (e.g., agency time and resources needed during transition to a managed behavioral healthcare system), implementation (e.g., data lag), agency issues (e.g., staff turnover), and client factors (e.g., motivation). Interventions took place during a time of Medicaid expansion and planning for statewide integration of mental health and substance use disorder treatment into a managed care model, which may have resulted in agencies not responding to the interventions. Moreover, incentives and alerts at the agency-level may not be effective when factors are at play beyond the agency's control.

Keywords

incentives; alerts; performance measures; outpatient treatment engagement; substance use disorder

1. Background

Financial incentives for quality improvement and timely performance feedback to treatment agencies are two potentially effective strategies for improving treatment quality. In the general medical field, evidence is generally weak for the effectiveness of incentive payments (Markovitz & Ryan, 2016) and is mixed for feedback (Ivers et al., 2012). To date, treatment for individuals with substance use disorders has been the focus of only a few empirical studies of these strategies. Therefore, we examined the impacts of these interventions by randomizing outpatient substance use treatment agencies in Washington State into four research arms that receive one or both of these interventions, or neither of them. Treatment agencies in the incentives arm could earn financial incentives for meeting a performance goal or for improving their own performance. Agencies in the alerts arm received weekly feedback on specific clients with regard to meeting the criteria on a performance measure. Agencies in the third arm received both interventions, and those in the control arm received neither. Expanding the focus of empirical work on incentives and feedback is essential in light of the need to improve quality of treatment for individuals with substance use disorders, as well as ongoing interest in payment reform at state and national levels.

1.1 Performance Based Contracting

Performance-based contracting is one approach to improving the quality of treatment, along with other initiatives such as workforce expansion, efforts at organizational change in treatment agencies (McCarty, Gustafson, Capoccia, & Cotter, 2009), system-level changes such as greater integration with general medical care (M T Stewart et al., 2017 under review) or increased reliance on managed care organizations. Performance-based contracting (PBC), in which payment is contingent upon meeting specified performance levels on a set of specified quality measures, is based on the logic that organizations respond rationally to incentives (Damberg et al., 2014; Damberg et al., 2007; Kondo et al., 2016; Markovitz & Ryan, 2016). Pay for performance (P4P), a general term for initiatives that are focused on quality improvement, has been adopted as an approach to quality improvement by private and public payers (CMS, 2011; James, 2012).

Recent reviews show mixed or relatively weak evidence that P4P improved performance in general medicine (Damberg et al., 2014; Eijkenaar, 2012; Markovitz & Ryan, 2016) and in substance use disorder treatment (Hull & Ritter, 2014), with a review of behavioral health studies suggesting that external incentives lead to positive outcomes (R. E. Stewart, Lareef, Hadley, & Mandell, 2017). Studies of implementation of incentive payments focused on substance use disorders have been focused on the public sector (Brucker & Stewart, 2011; Commons, McGuire, & Riordan, 1997; Lu, Albert Ma, & Yuan, 2003; Lu & Ma, 2006; Shen, 2003; M. T. Stewart, Horgan, Garnick, Ritter, & McLellan, 2013). In 2001, Delaware implemented performance based contracting (PBC) for outpatient treatment agencies based on performance targets established in contracts with the state agency (Haley, Dugosh, &

Lynch, 2011; McLellan, Kemp, Brooks, & Carise, 2008; M. T. Stewart et al., 2013), showing increased outpatient utilization (McLellan et al., 2008) as well as shorter waiting time for treatment, improved engagement and longer length of stay (M. T. Stewart et al., 2013). In 1992, Maine implemented PBC for substance use treatment (Commons et al., 1997; Lu et al., 2003; Lu & Ma, 2006; Shen, 2003) and in 2007 redesigned their approach (Brucker & Stewart, 2011). Recent analyses show the newer incentive contract did not have an effect on client access to care, nor was there any evidence of selection of clients more likely to meet the measures (M T Stewart et al., 2017 under review). Connecticut used an experimental design to assign programs to a control group or to incentives based on two measures: connection to a lower level of care after intensive residential treatment and readmission to detoxification (Tompkins et al., 2010), finding that the implementation was feasible but findings were mixed.

The lag in P4P initiatives focused on treatment for substance use disorders has been due, in part, to less consensus on performance measures and improvement strategies, less suitable state data, and lack of infrastructure in the public sector in some states to calculate quality measures to support P4P implementation for behavioral health (Bremer, Scholle, Keyser, Houtsinger, & Pincus, 2008). As noted in a review of P4P studies focusing on substance use disorder treatment programs (Hull & Ritter, 2014), the field lacks high quality randomized controlled trials that focus on organizational incentives.

Incentive payments may have differential impacts on specific groups of clients such as those in different age groups, racial/ethnic groups, or living in urban or rural settings. Some have warned that P4P may have the unintended consequence of worsening disparities if lower performing providers tend to care for poorer patients and have a large share of clients from racial/ethnic minority groups, and these providers receive lower incentive payments (Chien, Chin, Davis, & Casalino, 2007; Friedberg, Safran, Coltin, Dresser, & Schneider, 2010). Others have argued that P4P is a promising tool for reducing disparities in quality by socioeconomic status or race/ethnicity if incentive payment is designed to reward improvement as well as achievement (Blustein, Weissman, Ryan, Doran, & Hasnain-Wynia, 2011; Weissman et al., 2012). Because most of the work on differential impacts of incentives has been in general medicine, it is important to examine whether the use of P4P in behavioral health increases disparities.

Another possibility is that incentives are successful only in changing performance for agencies with a low baseline, and thus the potential to improve. Studies in the general medical sector show mixed results with one study reporting more improvement for the lowest baseline group (Greene, Hibbard, & Overton, 2015), but another study found that P4P did not result in any improvement even among lowest performers (Figueroa, Tsugawa, Zheng, Orav, & Jha, 2016).

1.2 Feedback through Reports of Performance, Alerts and Reminders

Another approach for quality improvement is delivery of electronic reminders to healthcare providers to indicate when a client is in danger of not receiving targeted services on a timely basis (Bryan & Boren, 2008; Garg et al., 2005). A Cochrane Collaborative review of controlled trials found that alerts and feedback can be effective in improving provider

performance, although the effects have generally been small to moderate and the effect may vary based on how the intervention is designed and delivered (Ivers et al., 2012). Electronic reminders have been shown to improve meeting recommended guidelines related to prescribing (Schedlbauer et al., 2009), prevention and screening procedures (Balas et al., 2000; Dexheimer, Talbot, Sanders, Rosenbloom, & Aronsky, 2008; Feldstein et al., 2006) and management of chronic conditions (Bryan & Boren, 2008). Moreover, more effective prompts and alerts are patient-specific, come from a well-respected leader and include evidence-based advice (Feldstein et al., 2006), provided before patient appointments (Balas et al., 2000; Dexheimer et al., 2008; Garg et al., 2005), and delivered at an appropriate point in the clinical workflow to make it easy for the provider to act on the information received (Feldstein et al., 2006). Graphical feedback regarding performance has also been shown to improve clinician protocol adherence (Andrzejewski, Kirby, Morral, & Iguchi, 2001).

1.3 Target Performance Measure

We focused on treatment engagement, a quality measure that has been included in the National Committee for Quality Assurance's (NCQA) set of HEDIS® measures (National Committee for Quality Assurance, 2017) and also adopted by the National Quality Forum (NQF)(National Quality Forum, 2012), as well as federal agencies such as the Centers for Medicare and Medicaid (CMS) and the U.S. Department of Veterans Affairs (VA). The NCQA definition combines initiation and engagement in a two-part measurement:

- **Initiation of AOD treatment:** Percentage of patients who initiate treatment through an inpatient alcohol or other drug (AOD) admission, outpatient visit, intensive outpatient service or partial hospitalization within 14 days of the diagnosis.
- **Engagement of AOD treatment:** Percentage of patients who initiated treatment and who had two or more additional services with a diagnosis of AOD within 30 days of the initiation visit.

In this study, we focus only on engagement and include only clients who are admitted to outpatient treatment in specialty settings (Garnick, Lee, Horgan, Acevedo, & Washington Circle Public Sector, 2009). Meeting the engagement criteria is associated with positive client outcomes such as reduced substance use and criminal justice involvement, and improved employment status among public-sector clients (Dunigan et al., 2014; Garnick et al., 2014; Garnick et al., 2007; Garnick et al., 2012) and improved outcomes among Veterans Administration clients (Harris, Humphreys, Bowe, Tiet, & Finney, 2010; Harris, Kivlahan, Bowe, Finney, & Humphreys, 2009).

1.4 Study Goals

In spite of mixed findings on the effectiveness of P4P, it remains a popular payment method with private and public payers as an approach to quality improvement (James, 2012). The evidence on the use of feedback and alerts has also resulted in modest improvements (Ivers et al., 2012). However, the work in both of these areas has been focused on general medicine and is relatively limited in behavioral health. As one of the first randomized studies of state incentive programs focused on substance use disorders, the overall goal of this study was to

explore the impact on performance of implementing two interventions in Washington State: financial incentives and a client-specific alert system. Focusing on the performance measure of engagement in outpatient treatment for substance use disorders, we addressed three questions:

- Does offering incentives or providing client-specific alerts to treatment agencies lead to increased likelihood of client treatment engagement?
- Do client-specific alerts in combination with incentives lead to an additional increase in client treatment engagement beyond that of incentives only or alerts only?
- Is there a differential impact of incentives only, alerts only, or incentives plus alerts on treatment engagement by subpopulations (e.g., age, racial/ethnic minorities and rural clients), or agencies with low baseline performance?

2. Methods

For this study, Washington State agencies providing publicly-funded outpatient (including intensive outpatient) substance use treatment in specialty settings were randomized into four study arms: incentives, alerts, incentives + alerts, and a control arm receiving no intervention. We used data from outpatient clients treated at these agencies during both pre- and post-periods to determine the impact of the interventions. The interventions were implemented simultaneously at agencies on October 1, 2013. Clients admitted to programs in the five quarters prior to the intervention date (July 1, 2012 to September 30, 2013) formed pre-intervention groups, and clients admitted from the intervention date through December 31, 2015 formed intervention groups. The study was approved by the Brandeis University and Washington State Institutional Review Boards.

2.1 Sampling Frame and Random Assignment

To prevent use of engagement rates with low reliability, prior to assigning to research arms, only treatment agencies with a minimum of 25 clients admitted during the pre-period were included, resulting in 151 agencies. Treatment agencies which were part of the same umbrella organization and in the same county would have similar management practices and may be in closer communication than independent agencies, so we randomized them as a group to the same arm. In addition, to create intervention arms which were as balanced as possible, we stratified our randomization by baseline performance level (three levels based on agency engagement rate during the pre-period), size (three levels based on the number of treatment admissions during the pre-period, and whether the treatment agencies were located in a rural area or not. Randomization resulted in 39 agencies assigned to the Incentives arm, 37 to the Alerts arm, 39 to the Incentives + Alerts arm, and 36 to the Control arm. The uneven number of agencies in the different arms resulted from the agencies under the same umbrella organization being randomized as a group, and several agencies closing.

2.2 Design and description of Interventions

2.2.1 Incentives—We based the amount of financial incentives that agencies were eligible to receive each quarter on earned points which took into account each agency's engagement

rate in that quarter compared to a target benchmark and improvement from its own performance baseline. This approach built on a point system approach initially developed by researchers in the Brandeis/Harvard NIDA Center (Tompkins et al., 2010) and implemented by CMS where hospitals are rewarded for how they perform on a set of specified measures and improve their performance relative to their baseline (CMS, 2007; James, 2012).

Agencies in the Incentives and Incentives + Alerts arms could earn points based on quarterly performance rates in two ways:

- *Achievement.* Achievement scores were assigned based on the position of an agency's engagement rate relative to two values, a threshold value, above which an agency would have a score greater than zero, and a benchmark value, above which an agency would receive the maximum possible score (i.e., 10). The minimum achievement threshold was set as the 50th percentile of agencies' engagement rate; the benchmark, was set as the 90th percentile of agencies' engagement rates. Points were then determined using equal intervals between the threshold and benchmark engagement rates. Using pre-intervention data, the achievement threshold was an engagement rate of 76% and the benchmark an engagement rate of 90%. Thus, only agencies which had a quarterly engagement rate of 76% or higher could get achievement points in that quarter.
- *Improvement.* For improvement points, the threshold was each agency's own baseline engagement rate, rather than the median baseline rate across all agencies. An agency earns an improvement score of 0 if its rate is below its own baseline value (i.e., no improvement), and it earns a score of 1 to 9 if its rate is in between (i.e., shows some improvement, but not up to benchmark yet).

Each agency's incentive payment was based on the higher of the two scores. The hope was that by designing the scoring as a combination of achievement and improvement, all agencies would remain motivated and continue in their efforts to earn incentive payments. Treatment agencies had to have at least 20 admissions in a quarter in order to be eligible for incentives in that quarter; otherwise, data from subsequent quarters was combined until reaching 20 admissions. Each quarter, an agency was awarded the maximum of its achievement or improvement points. The performance score was combined with the number of admissions in a quarter to determine if, and how much, each treatment agency would receive as incentives for the quarter. A payment table showed incentives based on the total number of points and the number of admissions in a quarter (Washington State DSHS, n.d.-b). Interventions ran for nine quarters, thus agencies in the Incentives and the incentives + alerts arms had nine opportunities to earn points and receive financial payments.

The following example illustrates how the points and incentives were applied. Outpatient Agency "A" had a baseline engagement rate of 40%. The achievement range was from 76% to 90% (threshold to benchmark), and their agency-specific improvement range was 40% to 90% (Agency A's baseline to the benchmark). In one quarter during the intervention, Agency A had an engagement rate of 54%. It did not earn achievement points because it did not reach the 76% minimum threshold. However, it earned 2 improvement points for improvement from 40% (their baseline engagement rate) to 54% (engagement rate in the

quarter). Based on the payment table (Washington State DSHS, n.d.-b), Agency A earned \$1,120 that quarter.

2.2.2 Alerts—Weekly alerts were sent by DBHR to agencies in the form of an Excel workbook with information about clients in the agency, on their deadline for meeting initiation or engagement criteria was approaching. The alerts showed the client name, the initiation and engagement deadlines, the number of days until the deadline, and the number of additional OP service days needed for the client to meet the criteria for initiation or engagement. While not the focus of the incentives, initiation is included because without meeting the initiation deadline, the client cannot meet the engagement criteria. The workbook also included graphs showing the treatment agency's quarterly performance rates. The file was sent weekly through secure email by DBHR to a staff member(s) designated by the agency. The accompanying email included tips for improving client engagement and provided links to a website with compilations of internet resources on how to increase client engagement.

2.2.3 Collaboration with agencies—Building on design features suggested by the literature, we sought and incorporated agency feedback about the design of the alerts and incentives, as well as the performance measures calculations and implementation, at four in-person meetings of the standing Washington Systems Improvement Workgroup (SIWG). The SIWG was comprised of treatment agency staff, county alcohol and drug coordinators, and DBHR staff who met quarterly to address systems issues, barriers and patient flow. Since not all agencies were part of the SIWG, we also administered two online surveys to all agencies, as well as an online survey to all agencies which had been assigned to receive alerts to obtain feedback.

2.2.4 Dissemination of Project Information to Agencies—Acknowledging that awareness and familiarity with the methods and goals of the study would be essential to its success, information about the study was widely disseminated. Agencies received letters describing the study and specifying the arm they were assigned to. They also had the opportunity to attend webinars about the project. Details of the study were presented at the annual Washington State's Co-Occurring Disorder Conference and published in an article that was included in the State's provider newsletter (Campbell, June 2013). We created a public website (Washington State DSHS, n.d.-a) where information about the study was posted, including webinar slides, agencies' randomization, and agencies' baseline and ongoing quarterly performance.

2.3 Sample

There were 28,657 outpatient admissions in the participating treatment agencies during the pre-intervention period and 49,526 during the intervention period. Six hundred and ninety-seven admissions in the pre-intervention period and 1,174 in the intervention period were excluded due to missing data. In addition, among the 17 outpatient treatment agencies which closed or had their contract with DBHR end during the intervention period, we excluded admissions ($N = 268$) that occurred in the six weeks before their last admission. Our rationale was that during such a disruptive time at an agency, clients would be less likely to receive

full attention and data on study effects would be biased. Also, clients with admissions close to the agency's contract end would not have enough time to engage. In all, our analyses included 27,960 admissions in the pre-intervention period (7/1/12 – 9/30/13) and 48,084 in the intervention period (10/1/13 – 12/31/15).

2.4 Data Source

Client treatment data were obtained from Washington's long established data system, the Treatment Activity Report Generation Tool (TARGET) for individuals receiving publicly funded substance use treatment. TARGET provides information on the dates and types of services, as well as information on demographics, employment, housing status, substance use history and current substance use.

2.5 Variables

Client-level treatment engagement was the key dependent variable for our analyses, indicators for the agency's assigned intervention arm were the key independent variables, and other client and agency characteristics served as covariates.

2.5.1 Treatment Engagement—Treatment engagement was defined as a client receiving at least one service within 14 days of outpatient treatment admission (treatment initiation), and two or more services within 30 days of initiation. Although it is possible for the same client to have been admitted more than once, we included all treatment admissions to make it easier for agencies to track their progress.

2.5.2 Client Covariates—We included the following client covariates: age, gender, race/ethnicity, past month use of specific substances, age at first use of any substance, education, marital status, employment status, homelessness, and criminal justice involved or referred. Most of these variables are based on clients' self-report and are collected at treatment admission. Criminal justice system involvement/referred was based on both self-report and/or participation in treatment funded by the state's Criminal Justice Treatment account. We also included an indicator variable noting whether the client lived in a rural area based on the client's zip code of residence. Using treatment data, we created variables indicating whether clients had, in the prior year, outpatient, residential, or detoxification services. Finally, using data from Washington's mental health treatment system linked to TARGET data, we controlled for whether clients had received any mental health services in the year prior to admission.

2.5.3 Agency covariates—Using aggregate client data, we created the following variables at the agency level for each quarter: agency size (based on number of admissions), percent of clients in the agency who reported using drugs other than alcohol or marijuana, percent of clients who had received mental health services in the prior year, and percent of clients who were involved in or referred by the criminal justice system. We also controlled for agency's participation in an online learning collaborative. The learning collaborative existed for 7 months and consisted of a series of webinars focused on treatment retention and using a web-based toolkit as a key resource (University of Washington Alcohol & Drug Abuse Institute, n.d.).

2.6 Analyses

We first conducted descriptive analyses of clients who were admitted to treatment at the participating agencies in the pre-intervention and intervention periods. We also employed a multilevel logistic regression model, with clients nested within agencies, to examine the effect of the interventions after adjusting for client and agency characteristics:

$$\text{Logit}(\text{engagement}_{ij}) = \beta_0 + \beta_1 \text{INT} + \beta_2 \text{ARM}_{ij} + \beta_3 \text{ARM}_{ij} * \text{INT} + \beta_4 \text{Client}_{ij} + \beta_5 \text{Agency}_j + u_j + r_{ij}.$$

In the model, *INT* is an indicator for the observation being during the intervention period; *ARM* is a set of three binary variables equal to 1 or 0 depending on whether the client is treated by each of our three interventions: financial incentive only, alerts only, and financial incentive plus alerts. *ARM*INT* is a set of interactions between the array *ARM* and *INT*. *Client* is the set of client level variables and *Agency* is the set of agency level variables. A statistically significant coefficient β_3 , associated with an *ARM*INT* term, would provide evidence of a differential impact of the *ARM* intervention on the likelihood of engagement.

To examine whether the interventions might exert a different impact on a subpopulation, we conducted stratified regressions based on client characteristics of age, race/ethnicity, and rural/urban client residence. We examined the agency characteristic of baseline treatment engagement rate, using the same three baseline performance categories used for randomization. We also tested the sensitivity of the results to examining intensive outpatient admissions separately.

2.7 Qualitative Focus Groups and Interviews

With the goal of gaining insights into the quantitative results, after the implementation period ended, we conducted a focus group and semi-structured phone interviews to gather feedback from agencies regarding the study interventions, how the interventions may have been improved, and perspectives on why the interventions did not have an impact on performance. These focus groups and interviews included representatives from outpatient as well as residential and detoxification agencies since the study included all three types of agencies.

The focus group was attended by 11 agency staff representing 9 agencies. Phone interviews were conducted with 17 staff members from 16 agencies, yielding a total of 28 individuals who provided feedback regarding the study and its interventions. All participants were compensated with a \$50 gift card. Responses were analyzed for common themes based on a conceptual model of factors that impact the effectiveness of incentives (Kondo et al., 2016) and using matrix analysis (Averill, 2002). Three authors analyzed the focus group transcript. After finding a high level of congruency on the major themes, one of the authors analyzed the interviews.

3. Results

3.1 Sample

Table 1 shows summary information on the demographic, substance use, and prior treatment characteristics for clients in the pre-intervention period and during the nine quarters of the interventions. Of the 76,044 outpatient clients in the sample, 32,344 were treated in regular outpatient and 43,700 were treated in intensive outpatient. The pre and post intervention samples of clients had similar characteristics.

3.2 Treatment agencies

A total of 150 agencies were assigned to one of the four research arms. Baseline engagement rates ranged from 13% to 100% (mean = 74; s.d. = 16). The mean number of admissions per agency ranged from 26 to 836 (mean = 154; s.d. = 155) in the pre-intervention period.

3.3 Unadjusted treatment engagement rates

Unadjusted treatment engagement rates by quarters are shown In Figure 1 for the five quarters before the study start (P1-P5) and for the nine quarters during the study intervention (Q1-Q5). The trends appear similar for all study arms, suggesting that there was probably little change in rates over time, and the interventions had no impact. The trends also appeared similar for regular outpatient and intensive outpatient when examined separately (data not shown).

3.4 Regression results

Table 2 shows estimates of the effects of intervention for the full sample of clients, controlling for client and facility covariates.

There were not significant main effects of the incentives only, the alerts only or incentives + alerts. Also, there was not an effect of time as indicated by the insignificant result for the intervention period variable. Finally, there was not any significant effect of the interaction of the time variable with any of the three intervention variables, indicating that the interventions did not have any significant effect on clients' engagement after outpatient admissions.

We also conducted stratified analyses of clients based on age, race/ethnicity, and urban/rural residence. These stratified results indicate that alerts or incentives, either alone or together, did not impact any effect on treatment engagement for clients in any subgroup. Analyses of clients from agencies with the lowest engagement rates at baseline (i.e., in the 33rd percentile or less) showed that even for clients treated in agencies with the most room for improvement, there was no significant impact from the interventions. We also conducted the same analysis for clients treated at agencies with the highest and mid-range engagement rates at baseline and found no significant impact of the interventions for their clients. Finally, in our sensitivity analysis, none of these results differed for clients receiving intensive outpatient treatment.

4. Discussion

Despite focusing on two promising approaches to improving the quality of treatment for substance use disorders, we found no evidence that incentives or alerts led to better client engagement in treatment. Moreover, these approaches did not improve treatment engagement for any specific category of clients by age, racial/ethnic group, or urban/rural residence, nor was there any improvement for regular outpatient as distinguished from intensive outpatient treatment. Contrary to other research findings (Greene et al., 2015), even among clients at agencies with lower baseline performance and thus ostensibly with the greatest room for improvement, we found no systematic improvement. Our findings, though, are consistent with the mixed evidence of the impact of incentive programs or alerts in general medical care (Damberg et al., 2014; Ivers et al., 2012; Markovitz & Ryan, 2016) and also consistent with the findings of a study on the impact of hospital value-based purchasing on quality of care (Figueroa et al., 2016). Regarding substance use treatment, our findings differ from results of an early intervention focused on substance use disorders in Delaware (M. T. Stewart et al., 2013), but are consistent with more recent findings in Maine (M. T. Stewart et al., under review).

Given the investment by Washington State's DBHR, treatment agencies, and researchers in designing and implementing the two interventions, these results are discouraging. However, the results may not be surprising given fundamental implementation challenges and contextual factors accompanying the project. Based on systematic reviews of the P4P literature as well as key informant interviews with P4P researchers, Kondo and colleagues developed a conceptual model of factors that impact the effectiveness of incentives (Kondo et al., 2016). In our focus group and post-study interviews with agency staff, we heard similar themes regarding relevant design features, implementation issues, treatment agency factors, and external factors affecting the effectiveness of our intervention in Washington State.

4.1 Effects of design features

Design features of incentive payment systems can impact effectiveness (Eijkenaar, 2012, 2013; Horgan, 2012; Rosenthal & Dudley, 2007), including the selection of the performance measure, the incentive design, nature of the target, and payment size.

Selection of Performance Measure—We selected treatment engagement as the target performance measure because it is nationally endorsed (Centers for Medicare and Medicaid Services, 2014; National Committee for Quality Assurance, 2016b), widely used (Harris, Bowe, Finney, & Humphreys, 2009; Harris et al., 2010), has evidence of being linked to outcomes, and is relatively simple to understand and track. Moreover, the optimal length of stay may vary by the phase of treatment (Proctor & Herschman, 2014), level of treatment, type of substance and patient characteristics. For years, Washington State's contracts with agencies had included a different performance measure, retention in outpatient treatment and many agencies may have viewed this retention measure as easier to achieve, since it required only one visit per month for three consecutive months and may have focused on the retention measure because it was in their contracts throughout the study period.

Incentive design—In designing the incentive payment structure, agencies were given the higher of two scores: how well the agency performed relative to all Washington outpatient agencies (called the ‘achievement score’) and how well the agency performed relative their baseline performance (called the ‘improvement score’). This approach recognizes that agencies which already have high engagement rates had less room and less need to improve, and at the same time it avoids the issue of low performing agencies not being sufficiently motivated. We selected the 90th percentile of all agency engagement rates as the benchmark because it was represented realistic level of excellence. This choice for a benchmark meant that 10% of agencies could earn a full achievement score of 10 merely by maintaining their baseline engagement rates. Moreover, given a relatively high average baseline engagement rate, 76% and potential “ceiling effects,” motivating additional clients to engage beyond the ones that already were engaging may have been more difficult, although those agencies in the bottom quartile of baseline performance certainly had room to improve. It is also possible that treatment engagement rates are difficult for agencies to improve. The NCQA HEDIS engagement measure has been flat or even declining in commercial and Medicaid managed care contexts (National Committee for Quality Assurance, 2016a).

Nature of incentives target—We aimed the incentives at treatment agencies rather than clients or clinicians because we wanted to give agencies the choice to use potential incentive payments in the manner they felt was most appropriate, including the possibility of distribution to staff or clients. In post-study focus group and interviews, however, some agency staff expressed a sense that incentivizing clients would have been more effective. Some agencies used client incentive approaches and reported they were effective in increasing treatment attendance and completion, approaches which are backed up by evidence from the literature on contingency management (Rash, Stitzer, & Weinstock, 2016).

Size of payment—The state allocated a substantial sum, \$1.5 million, towards incentive payments and in the hope of maximizing the effects of these payments, the study designed them to be distributed on a quarterly basis and to be based on agency size and performance. Post-study interviews indicated that while some agencies expressed that a larger incentive would have had more of an impact on performance, most felt that the size of the incentive did not make a difference and they just wanted to do the best for their clients. Some expressed that they found it odd to receive incentives for what they were supposed to be doing anyway, and some felt that the incentives implied they were not already trying their best already.

4.2 Implementation issues

Despite the robust data systems in Washington State and the participation of treatment agencies through the statewide Systems Improvement Workgroup (SIWG), certain unanticipated implementation issues arose.

Alert implementation—Because of concern that agencies would grow tired of receiving weekly alerts, we tracked the percent of alert e-mails that were opened in January and February 2014, finding that 54 % opened half or more of the alerts sent. However, looking at

the longer period from 10/1/2013 to 06/30/2014, 17% were opened in the alerts only arm and 35% were opened in the alerts and incentives arm, although the percent of alerts opened began higher and decreased over time. This suggests that the motivation to open the alerts was lower when there was not the possibility of financial incentives. However additional reasons may explain alerts not being opened or, when opened, not having an impact. Lags in when data was submitted to the TARGET system made the alerts less useful to agencies and several agencies reported that the data from the TARGET system were out of date by the time they received it.. The contractual requirements were that agencies submit data in a timely manner, but these requirements were not strictly enforced. For example, if an agency submits data on a monthly basis, weekly alerts with specific client information are not useful during the first part of each month. Other agencies found the alert system redundant because they had their own electronic health record system that had more current data. Although the alert design was vetted with the SIWG, some agencies found the alerts to be complex and reported that they would have appreciated technical assistance in order to use them. Communication of information from the state to agencies was less than ideal. Some agencies expressed the opinion that encrypted alerts requiring a passcode to open were not user friendly.

Incentive implementation—Our goal was to calculate incentive payments on a quarterly basis and to make payments as quickly as possible. This goal was thwarted at two key junctures. First, the need to allow 45 days to play out after the end of the quarter (from an initial treatment admission to possible initiation visit in 14 days to possible engagement visits in 30 days), as well as data lag in reporting to TARGET meant that final quarterly rates could not be calculated until seven weeks after a quarter ended. Second, the payments were administered through the counties that managed the outpatient agency contracts, which meant that there were additional delays. The State sent a letter to each agency in the incentives or incentives plus alerts arms on a quarterly basis reporting their baseline engagement rate, their quarterly engagement rate, their achievement and improvement points, and amount of the incentives earned, if any. To expedite feedback, in the second quarter of the intervention, we began sending an earlier letter based on preliminary data alerting agencies that were likely to get an incentive payment.

4.3 Treatment agency factors

Perhaps the most important element in the success of an incentive or alerts initiative is how agencies respond on an organizational level. The reaction we heard most often was that treatment agencies are stressed and have multiple competing demands for staff time and attention. Agencies expressed that staff turnover and limited staff time impacted the results of the interventions. In cases where the person designated to receive the alerts left the agency, the responsibility of opening the alerts may not have been transferred to another staff member. Agencies reported that it took time for staff to learn how to interpret and follow-up on the alerts, and with staff turnover this knowledge also may not have been transferred. In addition, staff time is limited to deal with the alerts and with following up with clients nearing the end of their time period for engagement.

Another reaction agencies expressed was a need for assistance with concrete steps to take to improve engagement rates. Some state agencies are doing this. For example, the Oklahoma Department of Mental Health and Substance Abuse Services (ODMHSAS) has implemented a technical assistance Data Integrity Review Team (DIRT) that visits agencies (University of Wisconsin, n.d.) to review their performance measures, provide one-on-one feedback, and make suggestions of steps to take in order to improve.

4.4 Treatment system issues

When we conceptualized this project and applied for funding in 2011, we did not foresee that Washington would be ramping up for significant changes in their treatment system, or anticipate the changes in the clients presenting for treatment. In addition, perennial issues outside of treatment agencies control impact their ability to engage clients. Until recently, the State provided physical health, mental health and SUD services to Medicaid enrollees through three separate systems. Following legislative actions in 2014, mental health and SUD services were integrated on April 1, 2016 and the plan is to fully integrate physical health and behavioral health by January 1, 2020. A number of agencies expressed that they have been in crisis mode with funding cuts and trying to keep up with changes, and have, therefore, had no time to focus on the interventions.

Our project coincided with a period of national transition as well. The Affordable Care Act (ACA) allows states to expand Medicaid to individuals who were previously ineligible, such as single adults without documented disabilities. However, agencies reported that Medicaid regulations are limiting. Outreach efforts that have the potential to get clients to come back for treatment (e.g., follow-up calls, going into the neighborhoods and knocking on doors) are not Medicaid reimbursable activities. Also, it is a “Medicaid rule” that in order to bill for otherwise allowed services, the person who calls the client for scheduling needs to be a CDP (chemical dependency professional). This requires clinical staff time that might otherwise be committed to other things.

At the same time, many agency staff reported that incentives are not sufficient in the face of client issues that cannot be controlled by the agency, such as client lack of motivation, barriers to treatment (e.g., lack of transportation and insurance coverage, or housing instability) and client contact information that gets outdated quickly so follow-up is difficult.

4.5 Limitations

The main limitations to our study are the lack of generalizability since the study was conducted only in one state, the lack of data on medication assisted treatment, and the analysis of data only from the public sector specialty treatment system. It could be that the impact of the interventions may be different in other settings or a more recent time period. For example, alerts may be more effective in settings with electronic medical records where they could be incorporated, for example, into real-time rather than weekly follow-up reminders. Moreover, it appears that the electronic reminders were not utilized as planned given technological and resource restraints. Thus, it is unclear whether the lack of an effect was due to the intervention not being effective or due to it not being used by some of the providers.

4.6 Conclusions and policy implications

In one of the first studies to use a randomized design and to focus on incentives and alerts as approaches to improving the quality of treatment of clients with substance use disorders, these interventions did not have any impact on performance. Design of a fair and effective pay for performance program is a complex undertaking (Eijkenaar, 2013; Horgan, 2012) and we found that even following many recommended design paths did not ensure success. Moreover, the large investment in implementation in terms of State personnel time and commitment of funds for incentive payments, and federal grant support was not sufficient to ensure improvement, raising the possibility that incentive payments and alerts, by themselves, may not be the most promising approach to improving performance for substance use disorders in public sector settings.

Due to disappointing or mixed results of P4P programs, much additional thought has been given to P4P design. Some view that neither the right incentives nor the right metrics have been designed (Jha, 2013), while others suggest that the right set of clients who would be most influenced by the intervention has not been identified (McKethan & Jha, 2014). Moreover, it may be that alerts or incentives could be successful in a context with improved data availability through electronic health records that could make data available in real time, thus facilitating providing alerts in real time and to specific clinicians or teams, offering incentives directly to clients, reaching out to clients through patient portals, or calculating incentives more quickly. Going forward, these improvements in technology are important areas to focus on along with questions such as how improvements, when achieved, can be sustained over time and whether use of performance measures has any unintended consequences.

Furthermore, it will be crucial to consider the influence of a changing healthcare landscape on the impacts of performance incentives or alerts. In Washington State in particular, treatment agencies were preoccupied with Medicaid expansion, planning for integration of mental health and SUD into regional managed behavioral health organizations (BHOs) which occurred in April 2016 and positioning themselves for full integration with medical services in 2020 under a managed care model. Moreover, at the time of the study, the treatment system may have lacked the capacity to respond to alert and incentives. Despite the current addition of resources through Medicaid expansion and new federal funding to respond to the opioid crisis (Pratt, 2017), it is critical to assess system capacity before launching any new quality improvement effort.

As states move forward in planning incentive-based approaches to improving the quality of treatment for clients with substance use disorders, these findings offer important lessons and support the need to plan carefully. One important lesson is that P4P be considered not only as a stand-alone approach, but rather as one element in a broader effort to drive quality improvement that might include multiple elements such as client-focused supports (e.g., client-focused incentives or recovery coaches) or strengthened support of treatment agencies (e.g., staff training or upgrades of health information systems). Moreover, states should evaluate their health system context in terms whether it will support improved quality in terms of agency- and state-level health information technology or the ability to attract sufficient numbers and skill-levels for an adequate treatment workforce. Perhaps only in the

context of a well-functioning treatment system, and in concert with a range approaches, will P4P emerge as a useful tool to drive quality improvement.

Acknowledgments

This research was supported by the National Institute on Drug Abuse (NIDA) of the National Institutes of Health (R01DA033468) and is a component project of the NIDA-supported Brandeis/Harvard Center to Improve System Performance of Substance Use Disorder Treatment (P30 DA035772). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or of Washington State. The authors appreciate the contributions of Can Du, Alice Huber, Eric Larson, Sharon Reif, Maureen Stewart, Katie Weaver-Randall, and Fritz Wrede.

References

- Andrzejewski ME, Kirby KC, Morral AR, Iguchi MY. Technology transfer through performance management: the effects of graphical feedback and positive reinforcement on drug treatment counselors' behavior. *Drug Alcohol Depend.* 2001; 63(2):179–186. [PubMed: 11376922]
- Averill JB. Matrix analysis as a complementary analytic strategy in qualitative inquiry. *Qual Health Res.* 2002; 12(6):855–866. DOI: 10.1177/104973230201200611 [PubMed: 12109729]
- Balas EA, Weingarten S, Garb CT, Blumenthal D, Boren SA, Brown GD. Improving preventive care by prompting physicians. *Arch Intern Med.* 2000; 160(3):301–308. [PubMed: 10668831]
- Blustein J, Weissman JS, Ryan AM, Doran T, Hasnain-Wynia R. Analysis raises questions on whether pay-for-performance in Medicaid can efficiently reduce racial and ethnic disparities. *Health Aff (Millwood).* 2011; 30(6):1165–1175. DOI: 10.1377/hlthaff.2010.1022 [PubMed: 21653971]
- Bremer RW, Scholle SH, Keyser D, Houtsinger JV, Pincus HA. Pay for performance in behavioral health. *Psychiatr Serv.* 2008; 59(12):1419–1429. DOI: 10.1176/appi.ps.59.12.1419 [PubMed: 19033169]
- Brucker DL, Stewart M. Performance-based contracting within a state substance abuse treatment system: a preliminary exploration of differences in client access and client outcomes. *The journal of behavioral health services & research.* 2011; 38(3):383–397. [PubMed: 21249461]
- Bryan C, Boren SA. The use and effectiveness of electronic clinical decision support tools in the ambulatory/primary care setting: a systematic review of the literature. *Inform Prim Care.* 2008; 16(2):79–91. [PubMed: 18713524]
- Campbell, K. Focus. Washington State Department of Social and Health Services, Division of Behavioral Health and Recovery; Jun. 2013 New Research Project Focusing on Improving Treatment Performance.
- Centers for Medicare and Medicaid Services. CLINICAL QUALITY MEASURES FOR 2014 EHR INCENTIVE PROGRAMS. 2014. Retrieved from https://www.cms.gov/regulations-and-guidance/legislation/ehrincentiveprograms/downloads/2014_ep_measurestable_june2013.pdf
- Chien AT, Chin MH, Davis AM, Casalino LP. Pay for performance, public reporting, and racial disparities in health care: how are programs being designed? *Med Care Res Rev.* 2007; 64(5 Suppl):283S–304S. DOI: 10.1177/1077558707305426 [PubMed: 17881629]
- CMS. Report to Congress: Plan to Implement a Medicare Hospital Value-Based Purchasing Program. 2007. Retrieved from U.S. Department of Health and Human Services
- CMS. [accessed May 11, 2016] Premier Hospital Quality Incentive Demonstration: Rewarding Superior Quality Care. 2011. Retrieved from <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/Downloads/HospitalPremierPressRelease-FactSheet.pdf>
- Commons M, McGuire TG, Riordan MH. Performance contracting for substance abuse treatment. *Health services research.* 1997; 32(5):631. [PubMed: 9402905]
- Damberg CL, Sorbero ME, Lovejoy SL, Mortsorf G, Raen L, Mandel D. Measuring Success in Health Care Value-Based Purchasing Programs: Summary and Recommendations. 2014 Retrieved from.

- Damberg CL, Sorbero ME, Mehrotra A, Teleki S, Lovejoy SL, Bradley L. An Environmental Scan of Pay for Performance in the Hospital Setting: Final Report. 2007 Retrieved from.
- Dexheimer JW, Talbot TR, Sanders DL, Rosenbloom ST, Aronsky D. Prompting clinicians about preventive care measures: a systematic review of randomized controlled trials. *J Am Med Inform Assoc.* 2008; 15(3):311–320. DOI: 10.1197/jamia.M2555 [PubMed: 18308989]
- Dunigan R, Acevedo A, Campbell K, Garnick DW, Horgan CM, Huber A, ... Ritter GA. Engagement in outpatient substance abuse treatment and employment outcomes. *J Behav Health Serv Res.* 2014; 41(1):20–36. DOI: 10.1007/s11414-013-9334-2 [PubMed: 23686216]
- Eijkenaar F. Pay for performance in health care: an international overview of initiatives. *Med Care Res Rev.* 2012; 69(3):251–276. DOI: 10.1177/1077558711432891 [PubMed: 22311954]
- Eijkenaar F. Key issues in the design of pay for performance programs. *The European Journal of Health Economics.* 2013; 14(1):117–131. [PubMed: 21882009]
- Feldstein A, Elmer PJ, Smith DH, Herson M, Orwoll E, Chen C, ... Swain MC. Electronic medical record reminder improves osteoporosis management after a fracture: a randomized, controlled trial. *J Am Geriatr Soc.* 2006; 54(3):450–457. DOI: 10.1111/j.1532-5415.2005.00618.x [PubMed: 16551312]
- Figuerola JF, Tsugawa Y, Zheng J, Orav EJ, Jha AK. Association between the Value-Based Purchasing pay for performance program and patient mortality in US hospitals: observational study. *BMJ.* 2016; 353:i2214.doi: 10.1136/bmj.i2214 [PubMed: 27160187]
- Friedberg MW, Safran DG, Coltin K, Dresser M, Schneider EC. Paying for performance in primary care: potential impact on practices and disparities. *Project Hope.* 2010; 29(5):926–932. doi:<http://dx.doi.org/10.1377/hlthaff.2009.0985>.
- Garg AX, Adhikari NK, McDonald H, Rosas-Arellano MP, Devereaux PJ, Beyene J, ... Haynes RB. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. *JAMA.* 2005; 293(10):1223–1238. DOI: 10.1001/jama.293.10.1223 [PubMed: 15755945]
- Garnick DW, Horgan CM, Acevedo A, Lee MT, Panas L, Ritter GA, ... Wright D. Criminal justice outcomes after engagement in outpatient substance abuse treatment. *J Subst Abuse Treat.* 2014; 46(3):295–305. DOI: 10.1016/j.jsat.2013.10.005 [PubMed: 24238717]
- Garnick DW, Horgan CM, Lee MT, Panas L, Ritter GA, Davis S, ... Reynolds M. Are Washington Circle performance measures associated with decreased criminal activity following treatment? *Journal of Substance Abuse Treatment.* 2007; 33(4):341–352. DOI: 10.1016/j.jsat.2007.03.002 [PubMed: 17524596]
- Garnick DW, Lee MT, Horgan CM, Acevedo A. Washington Circle Public Sector W. Adapting Washington Circle performance measures for public sector substance abuse treatment systems. *Journal of Substance Abuse Treatment.* 2009; 36(3):265–277. DOI: 10.1016/j.jsat.2008.06.008 [PubMed: 18722075]
- Garnick DW, Lee MT, O'Brien PL, Panas L, Ritter GA, Acevedo A, ... Godley MD. The Washington circle engagement performance measures' association with adolescent treatment outcomes. *Drug Alcohol Depend.* 2012; 124(3):250–258. DOI: 10.1016/j.drugalcdep.2012.01.011 [PubMed: 22364777]
- 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(4):673–680. DOI: 10.1377/hlthaff.2014.0998 [PubMed: 25847651]
- Haley SJ, Dugosh KL, Lynch KG. Performance contracting to engage detoxification-only patients into continued rehabilitation. *Journal of Substance Abuse Treatment.* 2011; 40(2):123–131. [PubMed: 21094591]
- Harris AH, Bowe T, Finney JW, Humphreys K. HEDIS initiation and engagement quality measures of substance use disorder care: impact of setting and health care specialty. *Popul Health Manag.* 2009; 12(4):191–196. DOI: 10.1089/pop.2008.0028 [PubMed: 19663621]
- Harris AH, Humphreys K, Bowe T, Tiet Q, Finney JW. Does meeting the HEDIS substance abuse treatment engagement criterion predict patient outcomes? *J Behav Health Serv Res.* 2010; 37(1): 25–39. DOI: 10.1007/s11414-008-9142-2 [PubMed: 18770044]

- Harris AH, Kivlahan DR, Bowe T, Finney JW, Humphreys K. Developing and validating process measures of health care quality: an application to alcohol use disorder treatment. *Med Care*. 2009; 47(12):1244–1250. DOI: 10.1097/MLR.0b013e3181b58882 [PubMed: 19786908]
- Horgan, C. Improving the Quality of Addiction Treatment through Provider Pay-for-Performance. Paper presented at the College on Problems of Drug Dependence 74th Annual Scientific Meeting: Contingency Management Working Group; Palm Springs, CA. 2012.
- Hull P, Ritter A. The potential role of pay-for-performance in alcohol and other drug treatment funding. 2014
- Ivers N, Jamtvedt G, Flottorp S, Young JM, Odgaard-Jensen J, French SD, ... Oxman AD. Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane Database Syst Rev*. 2012; 6:CD000259.doi: 10.1002/14651858.CD000259.pub3
- James J. Healthcare policy brief: Pay-for-performance. *Health Affairs*. 2012 Oct 11.
- Jha AK. Time to get serious about pay for performance. *JAMA*. 2013; 309(4):347–348. DOI: 10.1001/jama.2012.196646 [PubMed: 23340633]
- Kondo KK, Damberg CL, Mendelson A, Motu'apuaka M, Freeman M, O'Neil M, ... Kansagara D. Implementation Processes and Pay for Performance in Healthcare: A Systematic Review. *J Gen Intern Med*. 2016; 31(Suppl 1):61–69. DOI: 10.1007/s11606-015-3567-0 [PubMed: 26951276]
- Lu M, Albert Ma Ct, Yuan L. Risk selection and matching in performance-based contracting. *Health Economics*. 2003; 12(5):339–354. [PubMed: 12720252]
- Lu M, Ma C-tA. Financial incentives and gaming in alcohol treatment. *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*. 2006; 43(1):34–53.
- Markovitz AA, Ryan AM. Pay-for-Performance: Disappointing Results or Masked Heterogeneity? *Med Care Res Rev*. 2016; doi: 10.1177/1077558715619282
- McCarty D, Gustafson D, Capoccia VA, Cotter F. Improving care for the treatment of alcohol and drug disorders. *The journal of behavioral health services & research*. 2009; 36(1):52–60. [PubMed: 18259871]
- McKethan A, Jha AK. Designing smarter pay-for-performance programs. *JAMA*. 2014; 312(24):2617–2618. DOI: 10.1001/jama.2014.15398 [PubMed: 25375310]
- McLellan AT, Kemp J, Brooks A, Carise D. Improving public addiction treatment through performance contracting: The Delaware experiment. *Health policy*. 2008; 87(3):296–308. [PubMed: 18325621]
- National Committee for Quality Assurance. Initiation and Engagement of Alcohol and Other Drug Dependence Treatment. State of Health Care Quality Report. 2016a. Retrieved from <http://www.ncqa.org/report-cards/health-plans/state-of-health-care-quality/2016-table-of-contents>
- National Committee for Quality Assurance. Summary Table of Measures, Product Lines and Changes. 2016b. Retrieved from <http://www.ncqa.org/Portals/0/HEDISQM/HEDIS2016/HEDIS%202016%20List%20of%20Measures.pdf>
- National Committee for Quality Assurance. HEDIS 2017, Vol. 2, Summary Table of Measures, Product Lines, and Changes. 2017. Retrieved from <http://www.ncqa.org/Portals/0/HEDISQM/HEDIS2017/HEDIS%202017%20Volume%202%20List%20of%20Measures.pdf?ver=2016-06-27-135433-350>
- National Quality Forum. NQF Endorses Behavioral Health Measures. 2012. Retrieved from http://www.qualityforum.org/News_And_Resources/Press_Releases/2012/NQF_Endorses_Behavioral_Health_Measures.aspx
- Pratt, D. The Daily News. 19. Seattle, WA: 2017 Jan. Washington state to get funding from 21st Century Cures Act.
- Proctor SL, Herschman PL. The continuing care model of substance use treatment: what works, and when is “enough,” “enough?”. *Psychiatry J*. 2014; 2014:692423.doi: 10.1155/2014/692423 [PubMed: 24839597]
- Rash CJ, Stitzer M, Weinstock J. Contingency management: New directions and remaining challenges for an evidence-based intervention. *Journal of Substance Abuse Treatment*. 2016; doi: 10.1016/j.jsat.2016.09.008
- Rosenthal MB, Dudley RA. Pay-for-performance: will the latest payment trend improve care? *JAMA*. 2007; 297(7):740–744. DOI: 10.1001/jama.297.7.740 [PubMed: 17312294]

- Schedlbauer A, Prasad V, Mulvaney C, Phansalkar S, Stanton W, Bates DW, Avery AJ. What evidence supports the use of computerized alerts and prompts to improve clinicians' prescribing behavior? *J Am Med Inform Assoc.* 2009; 16(4):531–538. DOI: 10.1197/jamia.M2910 [PubMed: 19390110]
- Shen Y. Selection incentives in a performance-based contracting system. *Health services research.* 2003; 38(2):535–552. [PubMed: 12785560]
- Stewart MT, Horgan CM, Garnick DW, Ritter G, McLellan AT. Performance contracting and quality improvement in outpatient treatment: Effects on waiting time and length of stay. *Journal of Substance Abuse Treatment.* 2013; 44(1):27–33. DOI: 10.1016/j.jsat.2012.02.001 [PubMed: 22445031]
- Stewart MT, Horgan CM, Quinn A, Garnick DW, Reif S, Creedon TB, Merrick EL. The role of health plans in supporting behavioral health integration. 2017 under review.
- Stewart MT, Reif S, Mohr B, Nguyen A, Ritter G, Torres M, ... Horgan CM. Value-based Incentives in a Public Addiction Treatment System: Effects on Access and Selection. under review.
- Stewart RE, Lareef I, Hadley TR, Mandell DS. Can We Pay for Performance in Behavioral Health Care? *Psychiatr Serv.* 2017; 68(2):109–111. DOI: 10.1176/appi.ps.201600475 [PubMed: 27903140]
- Tompkins, C., Perloff, J., Horgan, C., Reif, S., Siembab, L., DiLeo, P. Feasibility of Financial Incentives for Substance Abuse Treatment Providers in Connecticut. Paper presented at the National Association of State Alcohol and Drug Abuse Directors (NASADAD) Annual Meeting; June 4, 2010; Norfolk, VA. 2010.
- University of Washington Alcohol & Drug Abuse Institute. Retention Toolkit. n.d. Retrieved from <http://adai.uw.edu/retentiontoolkit/>
- University of Wisconsin. Performance Management Implementation Story: Oklahoma's Journal From Being Good Data Collectors to STAR Data Users. n.d. Retrieved from https://www.niatx.net/toolkits/system/OK_STAR-SI_PerformanceManagement.pdf
- Washington State DSHS. Incentives and Alerts for Improving Substance Abuse Treatment in Washington State. n.d.-a. Retrieved from <https://www.dshs.wa.gov/bha/division-behavioral-health-and-recovery/incentives-and-alerts-improving-substance-abuse-treatment-washington-state>
- Washington State DSHS. A New Statewide Study On Improving Treatment Performance: Incentive Payment Tables. n.d.-b. Retrieved from https://www.dshs.wa.gov/sites/default/files/Washington%20Incentives%20Payment%20Tables%2004_21_15.pdf
- Weissman JS, Hasnain-Wynia R, Weinick RM, Kang R, Vogeli C, Iezzoni L, Landrum MB. Pay-for-performance programs to reduce racial/ethnic disparities: what might different designs achieve? *J Health Care Poor Underserved.* 2012; 23(1):144–160. DOI: 10.1353/hpu.2012.0030 [PubMed: 22643468]

Highlights

- No effect of incentives or weekly alerts on client engagement after outpatient admissions
- Explanations include: delivery system context (e.g., system capacity), implementation (e.g., data lag), agency issues (e.g., staff turnover), and client factors (e.g., motivation).
- Moreover, incentives and alerts at the agency-level may not be effective when factors are at play beyond the agency's control.

Figure 1.
Quarterly Engagement Rates

Table 1

Demographics of outpatient client admissions during pre-intervention (7/1/2012 – 9/30/13) and intervention (10/1/2013 – 12/31/2015) periods

	All Outpatient (N = 76,044)	
	Pre-Intervention (N = 27,960)	Intervention Period (N =48,084)
	N (%)	N (%)
<i>Female</i>	11,201 (40.1)	19,050 (39.6)
<i>Age</i>		
18–20	1,665 (6.0)	2,250 (4.7)
21–25	5,001 (17.9)	8,156 (17.0)
26–30	5,255 (18.8)	9,958 (20.7)
31–44	9,481 (33.9)	16,770 (34.9)
45+	6,558 (23.5)	10,950 (22.8)
<i>Race/ethnicity</i>		
White	18,236 (65.2)	31,440 (65.4)
Black	2,380 (8.5)	3,636 (7.6)
Latino	2,765 (9.9)	5,193 (10.8)
American Indian	3,132 (11.2)	4,935 (10.3)
Other race/ethnicity	1,447 (5.2)	2,880 (6.0)
<i>Education</i>		
< High school	8,207 (29.4)	13,587 (28.3)
High school degree/GED	15,786 (56.5)	26,890 (55.9)
> High school	2,298 (8.2)	4,567 (9.5)
Vocational	1,669 (6.0)	3,040 (6.3)
<i>Married</i>	6,335 (22.7)	11,326 (23.6)
<i>Homeless</i>	4,267 (15.3)	6,642 (13.8)
<i>Employed</i>	4,590 (16.4)	10,870 (22.6)
<i>Rural</i>	4,569 (16.3)	8,406 (17.5)
<i>Referred by/involved with criminal justice system</i>	16,143 (57.7)	28,226 (58.7)
<i>Substance(s) used in past month ^a</i>		
Alcohol		
none	10,847 (38.8)	20,443 (42.5)
1–3 times	4,846 (17.3)	8,213 (17.1)
4–12 times	3,890 (13.9)	6,345 (13.2)
13+ times	8,377 (30.0)	13,083 (27.2)
Marijuana		
none	16,263 (58.2)	27,825 (57.9)
1–3 times	2,468 (8.8)	3,935 (8.2)
4–12 times	2,006 (7.2)	3,499 (7.3)

	All Outpatient (N = 76,044)	
	Pre-Intervention (N = 27,960)	Intervention Period (N =48,084)
	N (%)	N (%)
13+ times	7,223 (25.8)	12,825 (26.7)
Cocaine	2,227 (8.0)	2,900 (6.0)
Opiates	7,519 (26.9)	13,463 (28.0)
Methamphetamines	8,083 (28.9)	16,617 (34.6)
Other Drugs	1,354 (4.8)	2,184 (4.5)
<i>Age of first use^b</i>		
10	3,508 (12.6)	5,718 (11.9)
11–14	10,417 (37.3)	17,334 (36.1)
15–17	8,514 (30.5)	14,409 (30.0)
18–20	2,996 (10.7)	5,729 (11.9)
21+	2,525 (9.0)	4,894 (10.2)
<i>Prior treatment</i>		
Prior OP	6,947 (24.9)	11,599 (24.1)
Prior residential	6,812 (24.4)	10,296 (21.4)
Prior detox	3,585 (12.8)	5,595 (11.6)
Prior mental health services	7,804 (27.9)	14,209 (29.6)

^a Substance was listed as a primary, secondary, or tertiary drug and frequency of use was one or more times in the past month or during the month of highest use in the last year.

^b Earliest age of first use of any of the substances reported as primary, secondary, or tertiary substance of abuse.

Table 2

Difference-in-Difference Logistic Regression Model of Impact of Interventions on Likelihood of Outpatient Client Engagement

	All Outpatient (N = 76,044)
	Estimates (95% CI)
Intervention Arm (Ref: Control Arm)	
Alerts Only	−0.095 (−0.471, 0.282)
Incentives Only	0.120 (−.254, 0.495)
Incentives and Alerts	0.021 (−.387, 0.429)
Intervention Period (Ref: Pre-Intervention)	
Intervention period	0.077 (−.079, 0.233)
Intervention Arm and Post Interaction	
Alerts Only X Intervention Period	−0.020 (−0.230, 0.191)
Incentives Only X Intervention Period	−0.096 (−0.306, 0.114)
Incentives & Alerts X Intervention Period	−0.021 (−0.224, 0.183)

Note: Controlling for client and facility covariates. Treatment engagement was defined as a client receiving at least one service within 14 days of outpatient treatment admission (treatment initiation), and two or more services within 30 days of initiation.