Published in final edited form as:

Drug Alcohol Depend. 2020 January 01; 206: 107735. doi:10.1016/j.drugalcdep.2019.107735.

Is It Feasible to Pay Specialty Substance Use Disorder Treatment Programs Based on Patient Outcomes?

Dominic Hodgkin^a, Deborah W. Garnick^a, Constance M. Horgan^a, Alisa B. Busch, MD^b,
Maureen T. Stewart^a, Sharon Reif^a

^aInstitute for Behavioral Health, Heller School for Social Policy and Management, Brandeis University

^bMcLean Hospital, and the Department of Health Care Policy, Harvard Medical School.

Abstract

Background: Some US payers are starting to vary payment to providers depending on patient outcomes, but this approach is rarely used in substance use disorder (SUD) treatment.

Purpose: We examine the feasibility of applying a pay-for-outcomes approach to SUD treatment.

Methods: We reviewed several relevant literatures: (1) economic theory papers that describe the conditions under which pay-for-outcomes is feasible in principle; (2) description of the key outcomes expected from SUD treatment, and the measures of these outcomes that are available in administrative data systems; and (3) reports on actual experiences of paying SUD treatment providers based on patient outcomes.

Results: The economics literature notes that when patient outcomes are strongly influenced by factors beyond provider control and when risk adjustment performs poorly, pay-for-outcomes will increase provider financial risk. This is relevant to SUD treatment. The literature on SUD outcome measurement shows disagreement on whether to include broader outcomes beyond abstinence from substance use. Good measures are available for some of these broader constructs, but the need for risk adjustment still brings many challenges. Results from two past payment experiments in SUD treatment reinforce some of the concerns raised in the more conceptual literature.

Conclusion: There are special challenges in applying pay-for-outcomes to SUD treatment, not all of which could be overcome by developing better measures. For SUD treatment it may be necessary to define outcomes more broadly than for general medical care, and to continue conditioning a sizeable portion of payment on process measures.

Corresponding author: Dominic Hodgkin, Heller School for Social Policy and Management, Brandeis University, Waltham, MA 02453, USA. <hodgkin@brandeis.edu>.

Contributors: All of the authors contributed to the conception and design of the paper. Dominic Hodgkin drafted the manuscript and all authors reviewed drafts of the manuscript and contributed edits and suggestions.

Approval statement: All authors have read and approved the final manuscript.

Conflict of Interest: No conflicts of interest to report.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

Pay-for-performance; pay-for-outcomes; provider payment; outcomes measurement; risk adjustment

1. Introduction

In recent years, efforts have grown to move the US health care system away from paying providers on a fee-for-service basis, and toward payment based on measures that address the value of the care provided rather than just the volume. These efforts have included Medicare's pay-for-performance (P4P) initiatives for hospitals and other providers, and the formation of accountable care organizations (ACOs), which agree to be rewarded partly based on their performance on previously agreed-upon quality measures. Substance use disorder (SUD) treatment has only occasionally been affected by these trends, with the inclusion of SUD treatment in some ACO arrangements (D'Aunno et al., 2015; Stuart et al., 2017). The development of some performance measures for SUD treatment (Garnick et al., 2012) has made its inclusion in these initiatives more feasible, and also allowed some SUD-specific P4P experiments by public payers (Haley et al., 2011; McLellan et al., 2008; Stewart MT et al., 2018; Stewart et al., 2013) and others (Garner et al., 2012; Klein et al., 2016; Vandrey et al., 2011). In the specific case of SUD treatment, various observers have criticized the current volume-based payment system for giving providers perverse incentives (Pating et al., 2012). For example, residential SUD providers who developed a successful treatment approach that resulted in fewer relapses would effectively be penalized under current payment approaches, because their admission volume would go down.

Within this increased focus on paying for value, purchasers such as states and insurers have become increasingly interested in trying to define value based on the clinical outcomes achieved, rather than solely on proxy measures such as the quality of care or use of evidence-based practices predictive of good outcomes. For example, since 2013 Medicare has been adjusting payments to private hospitals based on patient mortality, complications and healthcare-associated infections, and some private payers (Ryan et al., 2017) and Medicaid programs are following suit (Millwee et al., 2018). To date, this focus on paying for outcomes (P4O, a special case of P4P) appears to have been less common in SUD treatment than in general medical care (Center for Health Care Strategies, 2019; Hull and Ritter, 2014). However, that may change in the near future as payers gain confidence in P4O approaches based on their experience with general medical care. For example, a payer/provider coalition has launched pilots in several US states to create 'addiction recovery medical homes' that will be reimbursed partly based on clinical outcomes (Polak et al., 2018). Advocates for SUD treatment therefore need to become better-informed about P4O approaches. In particular, the field needs to evaluate whether the conditions truly exist for P4O to be applied to SUD treatment, whatever its merits for general medical care.

Given this context, the goal of this paper is to assess the feasibility and desirability of attempting to pay SUD specialty treatment programs based on patient outcomes. To accomplish this goal, we review several distinct literatures. First, we review economic theory

papers that describe the conditions under which P4O is feasible in principle. Second, we review the key outcomes expected from SUD treatment, and the measures of these outcomes that are available to payers. Third, we review the few actual experiences so far with paying SUD treatment providers based on patient outcomes.

2. Under What Conditions Is Pay-For-Outcomes Feasible?

The idea of ‘pay-for-outcomes’ is to identify which patient outcomes are most desirable and reward treatment providers based on the extent to which patients improve on those outcomes. This contrasts with the more common approach in most P4P programs of rewarding providers based on more proximal measures of structure, reporting or process (e.g. papers reviewed in (Horgan et al., 2018; Stewart et al., 2017)). This section discusses the relative merits of these different approaches, and the conditions which need to be met for P4O to be a feasible option. We draw on the ‘principal-agent’ literature in economics, which examines how a principal (e.g. a health care payer) can structure payment in a way that motivates an agent (e.g. provider) to deliver good outcomes, even when the principal cannot observe important aspects of the agent’s behavior (e.g. effort, quality of care) (Conrad, 2015; Sappington, 1991).

From the payer perspective, the most obvious advantage of rewarding providers based on patient outcomes is that it directs providers’ attention to what the payer (and the patient) values most. By contrast, rewarding providers for their performance on a process measure (e.g., treatment completion) may result in high performance on that measure, but may not improve results on the ultimate patient outcome measure (e.g., abstinence). In addition, in some contexts there may be multiple ways to improve outcomes (e.g., medication vs. psychotherapy), and providers are better informed than payers about the relative effectiveness of those different approaches, as well as in a better position than payers to tailor clinical recommendations to the needs of a specific patient. In such situations, rewarding providers based on specific process measures (e.g., medication was offered) means micro-managing their decisions about the appropriate choice of treatment. By contrast, paying for outcomes leaves the provider with more flexibility about how to achieve the ultimate shared goal: improved patient outcomes.

While the above analysis provides a useful framework, it misses some important features of health care, including presumably SUD treatment. The economic literature on procurement and contracting has identified some of these issues:

- **Provider control.** Process measures are easier for providers to control than patient outcomes. The latter, being distal rather than proximal, are often more strongly affected by factors beyond providers’ control, such as patient adherence, comorbidity, use of 12-step programs and mobile recovery-related applications, and work and home environments (Garnick et al., 2006). As a result, P4P imposes more financial risk on providers when it uses outcome rather than process measures, and providers may be rewarded or penalized for events beyond their control, which is unfair and weakens their incentives. This problem is also more serious if the outcomes are measured at a later time point, when the

provider has less influence over the patient, e.g. after discharge from a facility. In some health care contexts, payers try to address this issue by adjusting payments prospectively for measured patient characteristics (risk adjustment) (Conrad, 2015; Rosenthal et al., 2004) or by identifying intermediate outcomes that are less distal for providers (discussed below).

- **Attribution of provider effect.** When a patient sees multiple providers during the same episode, the payer must decide which provider gets the credit for the patient's ultimate outcome. (This can also apply to some process measures, e.g. treatment retention.) Traditional managed care often designates one provider as a gatekeeper who must approve referrals, and is held accountable for outcomes. More recently, accountable-care models have removed restrictions on referrals, and designate as accountable the physician who provides most of the care for a given patient.
- **Feasibility.** The administrative data systems used to measure provider performance are often good at capturing process measures, but less able to capture client outcomes, especially distal outcomes. For SUD treatment, accessing those data can raise confidentiality concerns related to 42 CFR Part 2, the federal privacy law that governs how and when SUD treatment information may be shared for health care purposes or research (US Government, 2017). Prior research documented provider concerns that the regulations caused legal confusion and inhibited communication and information sharing for clinical or research purposes (McCarty et al., 2017). In February 2017, the Substance Abuse and Mental Health Services Administration (SAMHSA) updated the Part 2 regulations, including definitions that would make administrative data more accessible, provided that adequate privacy and security protections for the data are in place (US Government, 2017); however, some challenges remain (Campbell et al., 2019). Responding to requests for additional clarity and support for care coordination while protecting privacy, in August 2019 SAMHSA proposed additional changes to 42 CFR Part 2 regulations; the comment period for these changes closed October 25, 2019.
- **Cherry-picking.** P4P programs motivate providers to select patients who are likely to perform better on whatever measure is being rewarded (Liao et al., 2018) (e.g. those with low-severity SUD, in a P4O program). This incentive could be stronger if the treatment program is rewarded based on outcomes, as it may perceive outcomes to be less controllable than process measures. In principle, adjusting payments for patient risk could address this, but the data available to payers are often too limited to fully correct for client differences.
- **Misreporting.** P4P, including pay-for-outcomes, may encourage misreporting of rewarded measures. E.g., if the payer rewards client abstinence as reported by the provider, then the provider is motivated to overstate abstinence. This problem is also present even when payers only reward process measures (Friedberg et al., 2018), but it could be more serious with P4O if payers find it harder to monitor outcomes than process measures.

Contextual requirements

Based on these considerations, one can ask what characteristics of the payment context would be required for outcome-based payment to be feasible. In a theoretical analysis, Chalkley and Khalil derived conclusions about the contexts in which outcome-based payment should be preferred to volume-based approaches (such as payments per diem or using diagnosis-related groups) (Chalkley and Khalil, 2005). We first describe these requirements before discussing whether they are likely to be met in the case of SUD treatment in the US.

First, outcome-based payments may be preferable in contexts where consumers are freer to exercise choice of whether or where to receive treatment. This is because in such contexts, consumers may be able to detect inappropriate treatment decisions and respond by refusing treatment or switching providers. Thus, consumers' ability to choose imposes some market discipline that limits providers' potential overreaction to the strong incentives present in outcomes-based payment. Chalkley and Khalil argue that consumers are freer to exercise choice for elective care but not emergency care, implying the former setting is better suited to outcomes-based payment (Chalkley and Khalil, 2005). One could also infer that outcomes-based payment is better suited to markets or types of care where there are multiple competing providers, since that facilitates consumer choice.

Second, Chalkley and Khalil conclude that outcomes-based payment would be more likely when the service providers have an intrinsic concern for patients. The rationale is that altruistic providers will respond less to any problematic incentives offered by outcomes-based payment, because distorting treatment decisions would reduce patient benefits and therefore also cause providers discomfort (Chalkley and Khalil, 2005).

Are these requirements likely to be met in the US SUD treatment sector? (See also Table 1). First, some patients have a choice of whether to receive treatment, but many do not, because they enter SUD treatment under court mandate, or under strong pressure from their employer, family or others (Polcin et al., 2012). Second, patients in many areas face little choice of where to receive SUD treatment. For example, nearly half the US population lives in counties with a shortage of opioid treatment programs (Dick et al., 2015), with particular problems in rural areas (Monnat and Rigg, 2018). Lack of choice is presumably greater for clients with more specialized needs (e.g. treatment for intravenous drug use while pregnant). Even where clients do face choice, one can question whether they have the knowledge to detect poor quality treatment or to choose providers based on quality. In fact, it may be the client's family or other parties that choose the provider. Third, altruism is certainly an important force for many individual clinicians, who entered the SUD treatment workforce to help people and (in many cases) in response to personal lived experiences. However, there is also evidence that SUD treatment programs do respond to payment incentives (Haley et al., 2011; Stewart et al., 2013), even sometimes including fraudulent activity such as overbilling, overuse of unnecessary tests, and payment of referral fees to sober home operators (Alvarez, 2017). These findings suggest that many treatment programs are strongly influenced by financial opportunities.

We conclude that SUD treatment in the US does not fully meet any of these contextual requirements for successful implementation of pay-for-outcomes.

Requirements for outcome measures

Even if the contextual requirements were met, P4O also requires the availability of accepted outcomes that can be validly and reliably measured. What attributes would those outcomes need to have? Propper and Wilson (Propper and Wilson, 2012) (summarizing (Smith, 2002)) proposed the following checklist: “outcomes are likely to be favoured when the nature of the outcome is relatively uncontested; it can be captured relatively easily in operational performance measures; when an indicator of the outcome can be secured reasonably soon after the intervention; when the outcome is readily attributable to clinical performance rather than external factors; and, when there is need for considerable clinical judgement as to the most appropriate intervention to offer.” In light of our earlier discussion, one could add two more conditions: that the outcome measure should not be easily misreported by providers, and that rewarding it should not encourage providers to shun certain clients. The latter condition would be met if a good risk adjustment methodology were available. After reviewing the actual measures available for SUD treatment, we will assess whether they meet these conditions.

3. What Patient Outcome Measures Are Available For SUD Treatment?

SUD treatment has multiple goals, complicating the question of how to measure and reward treatment outcomes. Below we list and discuss several of these desired goals.

Client's substance use

The most obvious goals of treatment relate to the client's use of problem substances, in particular the substance(s) that brought the client into treatment. Logically, payers and other stakeholders expect that treatment should result in the client ceasing the use of those substances (abstinence), or should at least substantially reduce the frequency and amount of use (harm reduction). On the other hand, the client may not agree with total abstinence as a goal, in which case it will be problematic to condition provider payment on abstinence. Nonetheless, the extent of substance use is one outcome that has been measured and rewarded in past P4O programs, and given stakeholder pressures, it is likely to be included in future programs.

For payers using substance use as a performance measure, one challenge is deciding at what point the client's substance use should be measured. In contrast to some other, generally acute conditions targeted for P4P, such as myocardial infarction, it can be hard to define exactly when a SUD treatment episode has ended, given the heterogeneity in the ways that patients begin and “end” treatment and that SUD requires ongoing care and monitoring, similar to other chronic conditions (U.S. Department of Health and Human Services (HHS) Office of the Surgeon General, 2018). If substance use is measured at the point when the client leaves a facility or completes a course of treatment, that may bias the measure toward finding treatment successful (even if relapse follows immediately thereafter). Measurement at discharge may also motivate providers to delay discharge for patients with continued

substance use, even if they are not benefiting from treatment (e.g. due to poor motivation). Payers could alternatively condition payment on the client's abstinence (or lower use) at some predetermined later point, e.g. 180 days after discharge. This would assess whether the treatment effects persisted beyond discharge. However, it also increases providers' financial risk for outcomes over which they have less control, particularly if the facility and client are no longer in contact after discharge. In addition, it can be costly and difficult to track down and interview SUD treatment clients after they have been discharged, whether this is to be done by the facility, the payer, or by independent evaluators. Supporting clients' recovery after discharge may require treatment programs to invest in additional services, but they often lack resources to do so (McLellan, 2011).

A related question is what should be the source of information about the client's substance use. Options include drug tests (for those substances where tests are informative), clinician report, or client self-report. Drug tests are expensive and may be difficult to obtain once the client is discharged. Further, they are inaccurate if the substance has metabolized out of a patient's system by the time the sample is taken—which occurs more frequently for substances with shorter half-lives (e.g., cocaine, alcohol) (Jaffee et al., 2008). Clinician report may be subject to gaming once it affects the provider organization's revenues. Self-report accords better with the current push toward greater use of patient-reported outcomes in health care (Pryor, 2017), can be more accurate than drug tests (Weiss et al., 1998), and is presumably more accurate in capturing the patient's own experience (Del Boca et al., 2016). However, patients may feel pressure to understate their actual substance use, resulting in biased measures, particularly in the absence of a therapeutic relationship.

Recovery status

In recent years there has been increased focus on recovery, rather than just abstinence, as a key goal of SUD treatment. A patient may be abstinent at the point of discharge, but if s/he is still experiencing strong cravings, or feels low self-efficacy in dealing with cravings, abstinence at discharge may not predict lasting success (Kadden and Litt, 2011; Tiffany and Wray, 2012). Beyond abstinence, many argue that treatment should be promoting recovery, which has been defined by SAMHSA as “a process of change through which individuals improve their health and wellness, live a satisfying self-directed life, and strive to reach their potential” (Substance Abuse and Mental Health Services Administration, n.d.). Various recovery-oriented outcome measures have been developed (Okrant, 2019), such as the Recovery Progression Measure, a 36-item survey based on cognitive behavioral therapy (Elison et al., 2016) and the Assessment of Recovery Capital, a 50-item survey addressing a variety of recovery pathways (Groshkova et al., 2013). However, they are not yet widely accepted by behavioral health stakeholders. Similarly, a recent review (Soper et al., 2017) noted that it has been difficult to standardize recovery concepts and link them objectively to payments.

In addition, there has been debate over when measurement of recovery should take place. McLellan and others have argued for periodic ‘concurrent recovery monitoring’ interviews with patients while they are still in outpatient treatment, to evaluate their progress toward symptom reduction and status improvement (McLellan et al., 2005). In fact, McLellan and

colleagues argue that improvement on those measures outside a controlled environment would constitute an outcome rather than a process measure (implying that its use for payment would constitute ‘paying for outcomes’).

Other client outcomes

Evaluations of SUD treatment often measure other outcomes considered to be desirable, even if they are not used in setting payment. Examples include medical consequences of use, and the client’s criminal justice activity, employment, and reconnection with family and community (Mant, 2001). Similar to the case of substance use outcomes, there are challenges about whether these outcomes are or should be available in administrative data accessible to payers, and if not, how to collect them, especially after discharge from treatment (Garnick et al., 2009). An expert panel has recommended augmenting provider EHRs with information on social determinants and functional outcomes (e.g. social isolation, stress), with a view to using them in value-based payment arrangements (Adler and Stead, 2015). US health plans and Medicaid programs are increasingly moving to collect data on social determinants of health, with encouragement from the Centers for Medicare and Medicaid Services (Artiga and Hinton, 2018; Threnhauser, 2018). Some are also experimenting with investment in those areas, as in North Carolina Medicaid’s recent Healthy Opportunities Pilots, or Massachusetts Medicaid’s move to risk-adjust capitation payments using area-level social determinants (Ash et al., 2017). However, even if patient-level information on these other outcomes was readily available, Soper *et al.* report that there are no currently endorsed outcome measures for domains like functional deficits in employment and education status, social connectedness, quality of life, and independent living skills (Soper et al., 2017). Additionally, these outcomes are affected by factors beyond provider control such as economic conditions, insurer interventions or policing strategies.

Performance measures

For an outcome to be included in P4O contracts, the measure needs to be operationalized in very clear terms: To what population should the measure apply, e.g., should some patients be excluded (the denominator)? What level of performance meets the measure, and under what conditions (the numerator)? In the US, a considerable research endeavor is aimed at developing and validating performance measures for health care in general, including SUD. Increasingly, entities that propose or develop performance measures have submitted them for endorsement to the National Quality Forum (NQF), a private non-profit that vets proposed measures in terms of criteria that include importance, scientific acceptability of measure properties, feasibility, and usability (Adirim et al., 2017). The NQF has not endorsed any outcome measures specific to SUD treatment, nor do any appear to be in the pipeline, although this has been noted as a gap (Garnick et al., 2012; Pincus et al., 2016).

Risk adjustment

Even if a payer had access to valid measures of all the outcomes of interest, paying based on those outcomes could be a problem if providers can profit by selecting only those patients likely to have good outcomes (biased selection). To avoid that possibility, P4P systems typically include an element of risk adjustment, which increases payment to providers who enroll a mix of patients with characteristics predictive of higher cost (Rosenthal et al., 2004).

We must therefore evaluate the usability of available risk adjustment techniques, not just outcome measures. In 2007, a review of available models concluded that those using diagnostic and sociodemographic information available from administrative data sets explained an average 7% of variance in mental health/SUD spending or utilization outcomes, whereas models using more detailed sources of data explained 23% (Hermann et al., 2007). More recently, Montz *et al.* concluded that for patients with mental disorders and SUD, the risk adjustment methodology used in state insurance exchanges would underpay health plans by 16% on average, because it under-predicted those patients' health costs (Montz et al., 2016). One issue is that variation in cost of care across SUD clients is likely to be strongly influenced by social determinants of health such as homelessness, family supportiveness or poverty, which are often not recorded in the data used to compute risk adjustment (Smith T et al., 2016.).

As an alternative to risk adjustment, another way to discourage cherry-picking behavior is to pay providers based on the improvement observed in each client from treatment entry to discharge, rather than simply outcome at discharge (Commons, 1997). This would better match the goals of payers and patients, but might pose feasibility issues and other challenges, e.g. ceiling effects.

The preceding discussion suggests that although some client outcome measures are available, they may not be strong enough to use as the main basis for paying providers. Measures appear to be more available and validated for measures of clients' substance use and abstinence than for some other important outcomes of treatment. However, paying providers solely based on substance use and abstinence might cause them to neglect aspects of treatment that mainly affect other outcomes, such as a broader definition of recovery, criminal involvement, mental health, or well-being.

4. Experience With Pay-For-Outcomes In SUD Treatment So Far

We only identified two experiments where P4O was applied to SUD treatment. These occurred in the US state of Maine (publicly funded treatment system, starting in 1992) and in eight small areas of the UK ("Payment by Results" pilots, 2012-14). Below we briefly review the characteristics of each experiment and findings regarding their effects.

Maine state Performance Based Contracting initiative (1990s)

In Maine, the state SUD agency provides grants (using federal block grant funding) to SUD treatment programs that maintain treatment slots open for state-funded clients, who typically do not have insurance. In 1992 the state agency introduced a performance-based contracting (PBC) system to pay for state-funded SUD treatment services (Table 2). The new system allocated grant funding across treatment programs in part based on their performance on a variety of outcome measures of efficiency, effectiveness, and service to special populations. The efficiency measures tracked service utilization. Of greater interest for this paper, the effectiveness measures tracked client outcomes on a variety of measures relating to substance use, employment, criminal involvement and relationships with spouse and family, as well as some process measures (Commons et al., 1997) (Table 3).

The contracts specified a minimum standard for each outcome indicator. Programs whose performance fell below standards in any of these areas were required to submit a corrective action plan to the state agency, while providers with good performance were rewarded with additional payments. The P4P design diverged from many commonly used designs in that it did not tell providers prospectively exactly what the rewards and penalties would be for different performance levels. This vagueness could have reduced the effectiveness of the incentives, unless providers were highly risk-averse about the potential to lose an unknown amount of funding.

The first major evaluation of this program used an uncontrolled pre-post design to assess the effect of PBC on the various performance measures, at the program level (Commons et al., 1997). For the client outcome measures ('effectiveness'), Commons *et al.* concluded that the PBC was associated with an increase in performance at those programs where a higher share of clients was state-funded. (Performance did not improve at programs where the share of state-funded clients was low.) Their measure of overall effectiveness performance was the proportion of monitored measures for which the program met the performance standards. They also ran separate analyses for each of the 15 effectiveness indicators (Table 3). These analyses identified a significant positive PBC effect for 9 of the indicators, and no significant effect for the other six indicators (Commons et al., 1997). A limitation of this study was the lack of a control group, although this was partly addressed by comparing treatment programs with higher versus lower proportions of state-funded clients. A second limitation was the reliance for outcome measures on client self-report as recorded by treatment program staff.

Subsequent research into the 1990s Maine PBC initiative using client-level data and more robust designs identified some unintended consequences, such as strategic misreporting of outcome data by programs (Lu, 1999; Lu and Ma, 2002, 2006), and selective avoidance of clients likely to look worse on the performance measures (Shen, 2003). These and other issues led the state to eventually discontinue its PBC initiative. Later, Maine's state agency instituted a second generation of performance contracts in 2007, with a variety of refinements to address lessons from the previous experience (Brucker and Stewart, 2011; Stewart MT et al., 2018). In particular, the new contract rewarded providers only based on process measures, not on patient outcomes, so it will not be further discussed here.

UK Drugs and Alcohol Recovery "Payment by Results" Pilot Program, 2012-14

In April 2012 the Department of Health in England introduced a pilot 'Payment by Results' (PbR) program to allow selected local public-sector health purchasing entities to pay providers of drug and alcohol misuse services based on 'recovery-focused' outcomes. The Department selected eight localities to be representative of all localities on demographic and other characteristics. In those localities, authorities started paying providers based on their clients' outcomes for drug use, crime, and health/wellbeing (Mason et al., 2015). During the design phase, one commentary noted various challenges facing the use of P4O for SUD, and advised proceeding "with caution" (Maynard et al., 2011).

The UK's National Institute of Health Research commissioned an evaluation of the experiment (Donmall et al., 2017). The study employed a difference-in-differences design,

with the comparison group consisting of the 141 English localities that did not join the experiment. The report (summarized in Table 3) found the intervention was positively associated with clients achieving abstinence, but negatively associated with successful completion of treatment within 6 months (for primary drug clients) and within 12 months of episode initiation (for both primary drug and primary alcohol clients). In addition, the intervention was negatively associated with primary drug clients presenting again for treatment within 12 months, among those who completed treatment within 6 months. This accorded with expectations, as one of the goals was to reduce readmissions. In addition, favorable effects were noted for acquisitive offending (theft and similar crimes; for drug clients only), but no significant association of the intervention with time to death. In several cases, these findings confirmed an earlier paper by the same team that reported on only the first year (Mason et al., 2015).

The evaluators offered explanations for these findings, based in part on interviews they conducted with providers, purchasers and others. For example, the penalty on readmissions (“re-presentations”) could have made agencies more cautious about discharging clients who seemed likely to seek readmission in the near future. If this led programs to retain clients longer this could help explain the decreased rate of successful treatment completion that was observed (Mason et al., 2015). Depending on the clinical appropriateness of retention, this actually could have been a good result.

The evaluation reported that all but one of the eight purchasing authorities planned to continue using P4O pay-for-outcomes, but with some modifications. These included: abandoning the use of offending as a payment-related outcome domain, shifting to a greater emphasis on process measures, and measuring readmissions over a six-month rather than a twelve-month period. One limitation of the evaluation was the fact that the areas implementing P4O were self-selected (volunteers). Another report criticized the initiative for diluting the focus on drug dependency by using so many outcome measures, and for allowing sites to exclude certain potentially higher-severity client groups (e.g. those receiving GP-prescribed methadone) from their measures (Centre for Social Justice, 2013).

5. Discussion

This paper has reviewed the requirements for a successful implementation of P4O for specialty SUD treatment programs, and suggested that at present they may not be fully met in the case of SUD treatment in the US. Future developments could alter this judgement, for example if outcome measures and risk adjustment methods undergo further improvement. However, meeting other requirements for successful P4O may be harder to achieve. For example, for the foreseeable future it is likely that the SUD treatment system will continue to have low levels of provider competition and limited influence on post-treatment outcomes. If P4O is introduced in contexts that do not meet the necessary requirements, various undesirable outcomes may result. These could include poorer performance on outcomes that are not rewarded; misreporting of the incentivized measures, and provider attempts to avoid patients who are less likely to perform well on the rewarded measures. Some of these challenges apply to general medical care too, but P4O in that sector has been

facilitated by greater agreement on treatment goals and wider diffusion of electronic health records, which have the potential to facilitate outcomes reporting.

This conclusion is supported by our review of the experience with the two major implementations of pay-for-outcomes to date, in Maine and the UK. Both initiatives rewarded a wide range of outcomes (not just reduced substance use), and attached rewards to process measures as well as to outcomes. Both initiatives resulted in improvements in abstinence rates, but they had mixed results on other rewarded outcomes. And both initiatives experienced deterioration in at least some important outcomes that were not being rewarded (e.g. validity of reporting in Maine; treatment retention in the UK). A recent report for the Australian government (using earlier results from the UK pilot), reached similar conclusions to ours: that explicitly linking funding to client outcomes would not be justified based on available evidence (Ritter et al., 2016).

It is also helpful to consider how far each experiment satisfied the requirements for successful operation of P4O, listed in Table 1. In each case, a substantial proportion of clients were in treatment under some coercion from the criminal justice system or elsewhere. In addition, patients did not apparently have a wide choice of providers in either case. In Maine, the state's rural character reduces choice of providers, while in the UK the localities' contracting practices appear to have reduced the number of providers offered to only one in 7 of the 8 areas (Mason et al., 2015). Regarding the degree of provider altruism, evidence on this is not available in either case. Regarding risk adjustment, Maine did not use this during the period evaluated (although for several measures it tracked improvement rather than final outcome), while the UK program used a specially developed tool which was later criticized as not very accurate for patients other than opioid users (Donmall et al., 2017). Thus, it appears that these programs did not fully satisfy a number of the requirements for successful operation of P4O, which may help explain the mixed results they achieved.

However, the above conclusions need not rule out a more limited use of outcome measures in combination with process measures and volume. This more mixed approach has also been the recommendation of various economic analyses for health care more generally (Eggleston, 2005), including Conrad's recommendation to put more weight on process measures than on outcomes, because they are more controllable by providers and provider organizations (Conrad, 2015). These considerations may be relevant if specialty SUD treatment comes under pressure from payers to join in P4O initiatives being applied to health care more generally. In that situation, SUD treatment providers may need to consider what tweaks could make P4O less risky, rather than rejecting it outright. In addition, the growing tendency of states and payers to collect data on social determinants of health could potentially lead to better measures of enrollee risk, allowing their use for risk adjustment in more refined P4O initiatives.

6. Conclusion

The theory and evidence reviewed in this paper should lead to caution regarding the extent to which patient outcomes should be used to pay specialty SUD treatment providers. Limited use could be acceptable if the payment system also continues to reward important process

measures, not just patient outcomes. In any future implementations of P4O, it will be important to build in an evaluation component and track providers' performance on both rewarded and unrewarded measures, including the extent to which each provider's case-mix is altered by selection activities. Finally, if the P4O program uses provider-reported data, payers should plan periodic audits to check the integrity of the measurement data being used for payment, including checking them against medical records (Lu and Ma, 2006).

Acknowledgment:

The authors thank Haiden Huskamp for comments on an earlier draft.

Role of Funding Source: Nothing declared. This research was supported by National Institute on Drug Abuse grants P30 DA035772 (the Brandeis-Harvard NIDA Center to Improve System Performance of Substance Use Disorder Treatment) and R01 DA033468 (PI: Reif). The funder had no involvement in study design; in the collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

Funding: This research was supported by National Institute on Drug Abuse grants P30 DA035772 (the Brandeis-Harvard NIDA Center to Improve System Performance of Substance Use Disorder Treatment) and R01 DA033468 (PI: Reif).

References

- Adirim T, Meade K, Mistry K, 2017 A New Era in Quality Measurement: The Development and Application of Quality Measures. *Pediatrics* 139(1).
- Adler NE, Stead WW, 2015 Patients in context--EHR capture of social and behavioral determinants of health. *N. Engl. J. Med* 372(8), 698–701. [PubMed: 25693009]
- Alvarez L, 2017 Haven for Recovering Addicts Now Profits From Their Relapses, *New York Times*.
- Artiga S, Hinton E, 2018 Beyond health care: the role of social determinants in promoting health and health equity.
- Ash AS, Mick EO, Ellis RP, Kiefe CI, Allison JJ, Clark MA, 2017 Social Determinants of Health in Managed Care Payment Formulas. *JAMA Intern Med* 177(10), 1424–1430. [PubMed: 28783811]
- Brucker DL, Stewart M, 2011 Performance-based contracting within a state substance abuse treatment system: a preliminary exploration of differences in client access and client outcomes. *J. Behav. Health Serv. Res* 38(3), 383–397. [PubMed: 21249461]
- Campbell ANC, McCarty D, Rieckmann T, McNeely J, Rotrosen J, Wu LT, Bart G, 2019 Interpretation and integration of the federal substance use privacy protection rule in integrated health systems: A qualitative analysis. *J. Subst. Abuse Treat* 97, 41–46. [PubMed: 30577898]
- Center for Health Care Strategies, 2019 Behavioral Health Provider Participation In Medicaid Value-Based Payment Models: An Environmental Scan And Policy Considerations. . <https://www.chcs.org/media/behavioral-health-provider-participation-in-medicaid-value-based-payment-models-an-environmental-scan-and-policy-considerations.pdf>.
- Centre for Social Justice, 2013 No Quick Fix: Exposing the depth of Britain's drugs and alcohol problem. <https://www.centreforsocialjustice.org.uk/library/no-quick-fix-exposing-depth-britains-drug-alcohol-problem> , Accessed 7/29/2018.
- Chalkley M, Khalil F, 2005 Third party purchasing of health services: patient choice and agency. *J. Health Econ* 24(6), 1132–1153. [PubMed: 15950301]
- Commons M, McGuire TG, Riordan MH, 1997 Performance contracting for substance abuse treatment. *Health Serv. Res* 32(5), 631–650. [PubMed: 9402905]
- Commons M, McGuire TG, 1997 Some economics of performance-based contracting for substance-abuse services, in: Egertson Joel A., Fox Daniel M., Leshner Alan I., al e. (Eds.), *Treating Alcohol Abusers Effectively*. Blackwell, Cambridge, MA.
- Conrad DA, 2015 The Theory of Value-Based Payment Incentives and Their Application to Health Care. *Health Serv. Res* 50 Suppl 2, 2057–2089. [PubMed: 26549041]

- D'Aunno T, Friedmann PD, Chen Q, Wilson DM, 2015 Integration of Substance Abuse Treatment Organizations into Accountable Care Organizations: Results from a National Survey. *J. Health Polit. Policy Law* 40(4), 797–819. [PubMed: 26124307]
- Del Boca FK, Darkes J, McRee B, 2016 Self-report assessments of psychoactive substance use and dependence. *The Oxford Handbook of Substance Use and Substance Use Disorders* 2, 430.
- Dick AW, Pacula RL, Gordon AJ, Sorbero M, Burns RM, Leslie D, Stein BD, 2015 Growth In Buprenorphine Waivers For Physicians Increased Potential Access To Opioid Agonist Treatment, 2002-11. *Health Aff. (Millwood)* 34(6), 1028–1034. [PubMed: 26056209]
- Donmall M, Sutton M, et al., 2017 Evaluation of the Drugs and Alcohol Recovery Payment by Results Pilot Programme: Final Report. http://research.bmh.manchester.ac.uk/epidemiology/NDEC/research/publications/PbRDR_Final_Report.pdf. Accessed 8/20/18.
- Eggleston K, 2005 Multitasking and mixed systems for provider payment. *J. Health Econ* 24(1), 211–223. [PubMed: 15617795]
- Elison S, Davies G, Ward J, 2016 Initial Development and Psychometric Properties of a New Measure of Substance Use Disorder "Recovery Progression": The Recovery Progression Measure (RPM). *Subst. Use Misuse* 51(9), 1195–1206. [PubMed: 27191336]
- Friedberg MW, Chen PG, Simmons M.L., Sherry T, Mendel P, Raaen L, Ryan J, Orr P, Vargo C, Carlasare L, Botts C, Blake K, 2018 Effects of Health Care Payment Models on Physician Practice in the United States, in: RAND Corporation (Ed.).
- Garner BR, Godley SH, Dennis ML, Hunter BD, Bair CM, Godley MD, 2012 Using pay for performance to improve treatment implementation for adolescent substance use disorders: results from a cluster randomized trial. *Arch. Pediatr. Adolesc. Med* 166(10), 938–944. [PubMed: 22893231]
- Garnick DW, Horgan CM, Acevedo A, McCorry F, Weisner C, 2012 Performance measures for substance use disorders--what research is needed? *Addict. Sci. Clin. Pract* 7, 18. [PubMed: 23186374]
- Garnick DW, Horgan CM, Chalk M, 2006 Performance measures for alcohol and other drug services. *Alcohol Res Health* 29(1), 19–26. [PubMed: 16767849]
- Garnick DW, Lee MT, Horgan CM, Acevedo A, 2009 Adapting Washington Circle performance measures for public sector substance abuse treatment systems. *J. Subst. Abuse Treat* 36(3), 265–277. [PubMed: 18722075]
- Groshkova T, Best D, White W, 2013 The Assessment of Recovery Capital: properties and psychometrics of a measure of addiction recovery strengths. *Drug Alcohol Rev* 32(2), 187–194. [PubMed: 22882622]
- Haley SJ, Dugosh KL, Lynch KG, 2011 Performance contracting to engage detoxification-only patients into continued rehabilitation. *J. Subst. Abuse Treat* 40(2), 123–131. [PubMed: 21094591]
- Hermann RC, Rollins CK, Chan JA, 2007 Risk-adjusting outcomes of mental health and substance-related care: a review of the literature. *Harv. Rev. Psychiatry* 15(2), 52–69. [PubMed: 17454175]
- Horgan C, Stewart M, Hodgkin D, Kritikos A, 2018 Rewarding value in substance use treatment: What are the unique challenges? , in: Brandeis University (Ed.). under review.
- Hull P, Ritter A, 2014 The potential role of pay-for-performance in alcohol and other drug treatment funding. University of New South Wales
- Jaffee WB, Trucco E, Teter C, Levy S, Weiss RD, 2008 Focus on alcohol & drug abuse: ensuring validity in urine drug testing. *Psychiatr. Serv* 59(2), 140–142. [PubMed: 18245154]
- Kadden RM, Litt MD, 2011 The role of self-efficacy in the treatment of substance use disorders. *Addict. Behav* 36(12), 1120–1126. [PubMed: 21849232]
- Klein AA, Lloyd KD, Asper LM, 2016 The feasibility of implementing a pay-for-performance program in the treatment of alcohol/drug addiction: Implementation and initial results. *Journal of Hospital Administration* 5(5), 1.
- Liao JM, Sommers BD, Navathe AS, 2018 Medicaid's Path to Value-Based Reform. *N. Engl. J. Med* 379(2), 105–108. [PubMed: 29996074]
- Lu M, 1999 Separating the true effect from gaming in incentive-based contracts in health care. *J Econ Management Strategy* 8(3), 383–431.

- Lu M, Ma CT, 2002 Consistency in performance evaluation reports and medical records. *J Ment Health Policy Econ* 5(4), 141–152. [PubMed: 14578548]
- Lu M, Ma CT, 2006 Financial incentives and gaming in alcohol treatment. *Inquiry* 43(1), 34–53. [PubMed: 16838817]
- Mant J, 2001 Process versus outcome indicators in the assessment of quality of health care. *Int. J. Qual. Health Care* 13(6), 475–480. [PubMed: 11769750]
- Mason T, Sutton M, Whittaker W, McSweeney T, Millar T, Donmall M, Jones A, Pierce M, 2015 The impact of paying treatment providers for outcomes: difference-in-differences analysis of the 'payment by results for drugs recovery' pilot. *Addiction* 110(7), 1120–1128. [PubMed: 26058447]
- Maynard A, Street A, Hunter R, 2011 Using 'payment by results' to fund the treatment of dependent drug users--proceed with care! *Addiction* 106(10), 1725–1729. [PubMed: 21793968]
- McCarty D, Rieckmann T, Baker RL, McConnell KJ, 2017 The Perceived Impact of 42 CFR Part 2 on Coordination and Integration of Care: A Qualitative Analysis. *Psychiatr. Serv* 68(3), 245–249. [PubMed: 27799017]
- McLellan AT, 2011 Considerations on performance contracting: a purchaser's perspective. *Addiction* 106(10), 1731–1732. [PubMed: 21917037]
- McLellan AT, Kemp J, Brooks A, Carise D, 2008 Improving public addiction treatment through performance contracting: the Delaware experiment. *Health Policy* 87(3), 296–308. [PubMed: 18325621]
- McLellan AT, McKay JR, Forman R, Cacciola J, Kemp J, 2005 Reconsidering the evaluation of addiction treatment: from retrospective follow-up to concurrent recovery monitoring. *Addiction* 100(4), 447–458. [PubMed: 15784059]
- Millwee B, Goldfield N, Turnipseed J, 2018 Achieving improved outcomes through value-based purchasing in one state. *Am. J. Med. Qual* 33(2), 162–171. [PubMed: 28750555]
- Monnat SM, Rigg KK, 2018 *The Opioid Crisis in Rural and Small Town America*. University of New Hampshire.
- Montz E, Layton T, Busch AB, Ellis RP, Rose S, McGuire TG, 2016 Risk-Adjustment Simulation: Plans May Have Incentives To Distort Mental Health And Substance Use Coverage. *Health Aff. (Millwood)* 35(6), 1022–1028. [PubMed: 27269018]
- Okrant E, 2019 *Developing a Patient-Reported Outcome Measure for Substance Use Disorder Recovery: A Person-Centered Approach to Monitoring Changes in Recovery Status* Brandeis University, The Heller School for Social Policy and Management.
- Pating DR, Miller MM, Goplerud E, Martin J, Ziedonis DM, 2012 New systems of care for substance use disorders: treatment, finance, and technology under health care reform. *Psychiatr. Clin. North Am* 35(2), 327–356. [PubMed: 22640759]
- Pincus HA, Scholle SH, Spaeth-Rublee B, Hepner KA, Brown J, 2016 Quality Measures For Mental Health And Substance Use: Gaps, Opportunities, And Challenges. *Health Aff. (Millwood)* 35(6), 1000–1008. [PubMed: 27269015]
- Polak AM, Smith D, Williams G, de Brantes F, 2018 *The Addiction Recovery Medical Home As An Alternative Payment Model*. Health Affairs Blog.
- Polcin DF, Korcha R, Greenfield TK, Bond J, Kerr W, 2012 Pressure to Reduce Drinking and Reasons for Seeking Treatment. *Contemp Drug Probl* 39(4), 687–714. [PubMed: 24431479]
- Propper C, Wilson D, 2012 *The use of performance measures in health care systems*, The Elgar companion to health economics, pp. 350–358.
- Pryor G, 2017 The data you've been missing — provided by patients. <https://www.athenahealth.com/insight/data-youve-been-missing-%E2%80%94-provided-patients>, accessed 8/17/2018.
- Ritter A, Hull P, Berends L, Chalmers J, Lancaster K, 2016 A conceptual schema for government purchasing arrangements for Australian alcohol and other drug treatment. *Addict. Behav* 60, 228–234. [PubMed: 27174218]
- Rosenthal MB, Fernandopulle R, Song HR, Landon B, 2004 Paying for quality: providers' incentives for quality improvement. *Health Aff. (Millwood)* 23(2), 127–141. [PubMed: 15046137]
- Ryan AM, Krinsky S, Maurer KA, Dimick JB, 2017 Changes in Hospital Quality Associated with Hospital Value-Based Purchasing. *N. Engl. J. Med* 376(24), 2358–2366. [PubMed: 28614675]

- Sappington DE, 1991 Incentives in principal-agent relationships. *J. Econ. Perspect* 5(2), 45–66.
- Shen Y, 2003 Selection incentives in a performance-based contracting system. *Health Serv. Res* 38(2), 535–552. [PubMed: 12785560]
- Smith P, 2002 Some Principles of Performance Measurement and Performance Improvement. University of York.
- Smith T, Myers R, Sederer L, J. B, 2016 How Value-Based Payment Arrangements Should Measure Behavioral Health. <https://www.healthaffairs.org/doi/10.1377/hblog20161129.057660/full/>, accessed 8/20/2018.
- Soper MH, Matulis R, Menschner C, 2017 Moving Toward Value-Based Payment for Medicaid Behavioral Health Services. <https://www.chcs.org/media/VBP-BH-Brief-061917.pdf>, accessed 8/20/2017.
- Stewart MT, Reif S, Mohr B, Nguyen A, Torres M, Davis MT, Ritter GA, Hodgkin D, CM H, 2018 Incentives in a Public Addiction Treatment System: Effects on Waiting time and Selection. *J. Subst. Abuse Treat*(95), 1–8.
- Stewart MT, Horgan CM, Garnick DW, Ritter G, McLellan AT, 2013 Performance contracting and quality improvement in outpatient treatment: effects on waiting time and length of stay. *J. Subst. Abuse Treat* 44(1), 27–33. [PubMed: 22445031]
- Stewart RE, Lareef I, Hadley TR, Mandell DS, 2017 Can We Pay for Performance in Behavioral Health Care? *Psychiatr. Serv* 68(2), 109–111. [PubMed: 27903140]
- Stuart EA, Barry CL, Donohue JM, Greenfield SF, Duckworth K, Song Z, Mechanic R, Kouri EM, Ebnesajjad C, Chernew ME, Huskamp HA, 2017 Effects of accountable care and payment reform on substance use disorder treatment: evidence from the initial 3 years of the alternative quality contract. *Addiction* 112(1), 124–133. [PubMed: 27517740]
- Substance Abuse and Mental Health Services Administration, n.d. SAMHSA's Working Definition of Recovery. <https://store.samhsa.gov/shin/content/PEP12-RECDEF/PEP12-RECDEF.pdf>, accessed 8/17/2018.
- Threnhauser SC, 2018 What Are Payers, Health Plans, & ACOs Doing About Social Determinants? .
- Tiffany ST, Wray JM, 2012 The clinical significance of drug craving. *Ann. N. Y. Acad. Sci* 1248, 1–17. [PubMed: 22172057]
- U.S. Department of Health and Human Services (HHS) Office of the Surgeon General, 2018 Facing Addiction in America: The Surgeon General's Spotlight on Opioids. Washington, DC.
- US Government, 2017 Confidentiality of Substance Use Disorder Patient Records. 82 Fed. Reg. 6052 (to be codified at CFR Part 2).
- Vandrey R, Stitzer ML, Acquavita SP, Quinn-Stabile P, 2011 Pay-for-performance in a community substance abuse clinic. *J. Subst. Abuse Treat* 41(2), 193–200. [PubMed: 21489739]
- Weiss RD, Najavits LM, Greenfield SF, Soto JA, Shaw SR, Wyner D, 1998 Validity of substance use self-reports in dually diagnosed outpatients. *Am. J. Psychiatry* 155(1), 127–128. [PubMed: 9433351]

Highlights

- Some insurers are paying general medical providers based partly on patient outcomes
- We reviewed literature on pay-for-outcomes: theory, measures, and application to SUD.
- There are special challenges in applying pay-for-outcomes to SUD treatment
- For SUD treatment it may be necessary to continue paying partly on process measures.

Table 1.

Requirements for successful use of pay-for-outcomes

Requirement	Requirement met for SUD treatment?	Comments
<u>Requirements for the treatment context</u>		
Patients have a choice of whether to receive treatment	Somewhat	Many patients enter SUD treatment under court mandate, or pressure from employer/family/ others
Patients have a choice of providers (competition)	Somewhat	Choice is present in some areas, absent in many
Altruistic orientation among providers	Somewhat	Some altruism present, but also some profit-seeking behavior
Accurate risk-adjustment methods	No	Available risk-adjustment methods are not very accurate
<u>Requirements for outcomes measures</u>		
Consensus about desired outcomes	No	Disagreement about relative importance of abstinence/use, crime and social outcomes, and concurrent outcomes (during treatment)
Outcomes measurable soon after the intervention	Not easily	Only with special data collection efforts
Outcome is readily attributable to clinical performance rather than external factors	Not necessarily	Particularly a problem for longer-term outcomes
Selection of the most appropriate intervention requires considerable clinical judgement	Yes	Choice of intervention is affected by patient characteristics, severity etc.
The outcome measure should not be easily manipulated by providers	Depends on how collected	Less manipulable if outcomes data are collected by independent evaluators (but that is costly)
The outcome measure should not encourage providers to shun certain clients	No	Until risk-adjustment methods improve, providers paid based on outcomes would benefit by shunning certain patients (poor prognosis)

Table 2:

Design and evaluation of pay-for-outcomes initiatives in Maine and the UK

	Commons 1997	Mason 2015
A. PROGRAM DESIGN		
Setting	Maine public treatment system, USA	8 pilot localities in England
Time period	1989-94	2011-13
Units receiving reward	Specialty SUD treatment programs	Specialty SUD treatment programs
Units paying reward	State substance abuse authority	“Drug Action Teams” that pay providers (“commissioning entities”)
What triggers reward	Program must exceed preset % of clients achieving the outcome	Varies by locality
Form of reward	Not explicit until 1994. Poor-performers had to meet state officials re how to improve. Some good performers were awarded additional federal block grant funds	Part of provider’s income was at risk for performance on outcomes, with the share varying across areas (from 5% to 100%).
Proportion of providers’ revenue at risk	Not reported; probably varied	Varied from 5% to 100%, depending on decision by each locality. Localities also chose how much weight to put on the different measures.
Risk adjustment	None. But many measures were computed as change from entry to discharge	Budgets were adjusted for client characteristics based on a specially developed model (the ‘complexity tool’)
B. EVALUATION		
Unit of analysis	Program-quarter	Admission
Patient outcomes evaluated	% of the effectiveness standards that the program met, in quarter	% of all drug treatment episodes resulting in ‘Successful completion’, including being judged by a clinician to be free of dependency from the drug for which the individual was being treated, and in addition not using either heroin or crack cocaine
Period studied	1989-1994	2010-2013
Study design	Pre-post with no comparison group	Pre-post with comparison group
Experimental group	All publicly funded programs in Maine were subject to the P4P initiative	The 8 areas in England that successfully applied to take part in this pilot program
Comparison group	None – but the model tested whether P4P had a stronger effect when the payer’s share of volume was larger (dependence)	The other 140 localities in England
Papers reporting on evaluation	Commons et al., 1997	Mason et al., 2015; Donmall et al., 2017

Source: adapted from information in Commons et al. 1997, Mason et al. 2015 and Donmall et al. 2017.

Table 3:

Impact of pay-for-outcomes initiatives in Maine and the UK

	Maine PBC		UK PBR		
	Rewarded?	Impact of initiative	Rewarded?	Impact of initiative	
				Drug clients	Alcohol clients
<u>A. CLIENT OUTCOMES</u>					
1. Substance use					
Abstinence:	Y	Favorable	Y	Favorable	No change
Reduction of use of primary substance abuse problem	Y	Favorable	Y	Not reported	Not reported
Treatment completion free from drugs/alcohol			Y	Unfavorable	Unfavorable
2. Employment					
Maintaining employment	Y	No change			
Employment improvement	Y	No change			
Employability	Y	No change			
Reduction in number of problems with employer	Y	Favorable			
Reduction in absenteeism	Y	No change			
3. Crime					
Not arrested for any offense	Y	Favorable			
Not arrested for OUI offense during treatment	Y	Favorable			
Acquisitive offending (e.g. theft)			y	Favorable	No change
4. Relationships					
Reduction of problems with other family members	Y	Favorable			
Reduction of problems with spouse/significant other	Y	Favorable			
<u>5. Other</u>					
Participation in self-help during treatment	Y	Favorable			
Housing problems			Y	No change	Not reported
Mortality			Y	No change	No change
Health and well-being			Y	Not reported	Not reported
<u>B. PROCESS MEASURES</u>					
Referral in continuum of care	Y	Favorable			
Referral to self-help	Y	No change			
Time in treatment	y	No change			
No re-presentation in either the treatment system or in the criminal justice system			Y	Favorable	No change
Unplanned discharge from treatment			Y	Unfavorable	Unfavorable
Treatment retention			Y	Favorable	Favorable
<u>C. OTHER</u>					
Selective misreporting of outcome data	y (implicitly)	Unfavorable			
Program-level severity mix	y (implicitly)	Unfavorable			

Source: adapted from information in papers about the interventions in Maine (Commons et al., 1997; Lu, 1999; Lu and Ma, 2002, 2006; Shen, 2003) and England (Donmall et al., 2017; Mason et al., 2015).

Notes:

1. Abstinence was measured 30 days prior to discharge in Maine, and at discharge in the UK.
2. In Maine, most outcomes are measured at discharge, and changes were measured between admission and discharge. Source: Commons et al., 1997.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript