Title of Study:

Individualizing Incentives for Alcohol in the Severely Mentally Ill

NCT number: NCT03481049

Date of Document: 6/27/2025

INTRODUCTION

We thank the reviewers for their enthusiastic and thoughtful feedback on the initial application (impact score = 37, 26th percentile). We respond to their concerns below and place revised text in **[**bold brackets.**]**

Unclear Power Estimates, Lack of Pilot Data. To address these concerns we removed the combined Shaping+High-Magnitude contingency management (CM) condition. We now propose a 3-group design to test whether the 2 CM adaptations (Shaping CM and High-Magnitude CM) outperform Usual CM and how they compare to one another in terms of clinical and cost outcomes. This change increases our statistical power, enables detailed power estimates based on previous research (Section 4.G.3), and permits better integration of the Addictions Neuroclinical Assessment (ANA) framework with our approach. We also now detail the study premise in Section 2.C. **Integration of ANA.** We now present hypotheses (Specific Aims & Section 2.A.6), background information (Section 2.A.6), and data analyses (Section 4.G.2) for our investigation of the relationship between ANA variables and clinical outcomes across the full study sample and in each adaptation of CM.

Each Study Component is Not Novel Alone. We rewrote Section 3 (Innovation). The revised proposal is highly innovative because our anticipated findings, combined with results from our initial funding period, will allow us to develop a personalized CM intervention for alcohol use disorders in which patients can be matched to CM conditions based on ethyl glucuronide (EtG) levels and ANA variables. Ours will also be the first CM study on alcohol use disorders to adapt CM for non-responders. Few prior studies have adapted CM; all have methodological flaws, and none addressed alcohol. Nor have any studies investigated whether adapted treatments for co-occurring disorders improve outcomes in non-responders. Given the poor outcomes and high costs of heavy drinking with co-occurring serious mental illness (SMI), it is vital to understand how treatment efficacy can be improved. Our economic analysis is also novel; it will be the first to address a CM intervention for alcohol use and it includes a comprehensive set of cost and benefit outcomes. Some outcomes, such as heavy drinking days, have not been studied previously. Our focus on the ANA is also highly innovative, because no previous CM study has used a theoretical-based framework to investigate predictors of treatment response, and no alcohol treatment study has used the ANA to predict outcomes. Thus, considering our strong study premise (Section 2.C) and improved rigor, this innovative proposal is likely to have high impact.

Outcomes Not Published. A manuscript summarizing results of the trial conducted in the initial funding period was accepted for publication in the *American Journal of Psychiatry* (Section 2.B & Appendix); these results were also presented at 3 national conferences. Three other manuscripts based on study outcome data are under review, including one on the link between pre-randomization EtG levels and CM outcomes. As noted in the Publication List, this grant contributed to the support of 18 papers and 21 presentations.

How Will Use of TAU be Assessed and Analyzed? We will monitor all units of non-study healthcare utilization, including substance use and mental health treatment as usual (TAU), with the Non-Study Resources Form (Section 4.F.6), which is commonly used to document TAU services in cost-effectiveness research.[1, 2] We will also assess use of TAU through administrative data documenting all TAU encounters. We will analyze whether increased use of TAU is associated with improved outcomes (Section 4.G.2). **How will Real World Treatment Costs be Estimated?** The purpose of our economic evaluation is to inform "real world" healthcare decisions. Therefore, we will employ a resource costing method to estimate patient-level costs. This involves determining a price weight for each resource unit consumed and multiplying price weights by units of service.[3-5] Per-person costs are then computed by multiplying units of service by the respective unit price and summing the values. The resources required to implement the intervention in non-study environments will be valued by using price weights that reflect specified stakeholder perspectives – the payer of healthcare services, the treatment provider, and society (Section 4.G.2). **How will Clinical Significance be Determined?** We used published thresholds for determining clinical significance based on Miller & Manuel (2008): 12% point differences for dichotomous variables and 50% point differences for continuous measures as the basis for our power analyses.[6] Results of previous CM studies are consistent with these thresholds. Our economic analyses will include time abstinent and heavy drinking days averted because these measures are clinically significant to consumers, clinicians, payers, and policymakers.

Will Mental Health Clinicians Implement CM? We now include data from our previous work demonstrating that CM is an acceptable approach to mental health clinicians (Section 2.A.3). Currently the primary barrier to CM implementation is cost.[7, 8] **Concerns about Recruitment.** We offer a more detailed justification for our recruitment projections and additional strategies to improve recruitment (Sections 4.C.1, 4.C.2). **No Evidence for EtG Point-of-Care Test.** We apologize for the confusion. We will use an EtG immunoassay, which requires an onsite benchtop analyzer, not point-of-care dipcards. This is necessary because point-of-care tests have a cutoff (300 ng/mL) that is insufficient to detect light drinking. Dipcards with lower cutoffs will be available in the near future, increasing CM feasibility. **Drinking Goals Cannot be Imposed.** We agree and will investigate whether drinking goals predict treatment outcomes by using an established measure (Section 4.F.7).[9]

Other Issues are addressed in the application, including **How Will Ability to Engage in Procedures be Assessed?** (Section 4.D.3), **Managing Suicidal and Homicidal Risk** (Human Subjects), **Several Personnel Not Named** (Budget Justification), **Lifetime Prevalence Overstates Annual Prevalence** (Section 2.A.1).

1. SPECIFIC AIMS

We propose to determine whether modifications to contingency management (CM) can improve outcomes and reduce costs in heavy drinkers with serious mental illness (SMI). Up to 46% of adults with SMI experience an alcohol use disorder in their lifetimes.[10] Alcohol use is associated with high rates of homelessness, hospitalization, HIV infection, cigarette smoking, and drug use in this group.[11-16] CM offers patients rewards for abstinence and is a promising addiction treatment for adults with SMI.[17-19] However, research on CM for alcohol use disorders (AUDs) has been limited by the lack of a feasible biomarker to verify abstinence.

In the initial funding period we found that the alcohol biomarker ethyl glucuronide (EtG) could detect drinking for up to 5 days in the context of a randomized trial of CM.[20-24] **[**In an article in the *American Journal of Psychiatry***]** we report that CM participants were **3 times more likely to submit alcohol-negative EtG tests** than controls.[25] They also reduced self-reported alcohol use and heavy drinking. The efficacy of CM was associated with EtG levels during a 4-week pre-randomization period (**Table 1**). Participants with mean pre-randomization EtG < 500 ng/mL (i.e., light drinking) attained **2.5 times longer duration of continuous abstinence during CM** relative to those with mean EtG > 499 ng/mL (i.e., recent heavy drinking). This difference represents 2 additional weeks of abstinence. In fact, heavy drinkers attained a duration of abstinence equal to controls, suggesting they did not respond to CM.

Table 1. Mean number of consecutive alcohol-negative samples

	CM M (SD)	**Control** M (SD)
EtG < 500 ng/mL	11.6 (10.7)	4.9 (3.5)
EtG > 499 ng/mL	4.6 (6.9)	3.7 (7.3)

Two potential CM adaptations might improve outcomes for heavy drinkers. One is **Shaping CM**, which reinforces reductions in use before requiring abstinence and is associated with positive outcomes in people with severe cocaine or nicotine addiction.[26-29] The other is **High-Magnitude CM**, which increases the magnitude of reinforcers for abstinence and is associated with reduced drug use in those who do not respond to a usual CM intervention or submit a drug-positive urine test prior to CM.[30-32] While both strategies are associated with increased abstinence in extant studies, no randomized trial has 1) compared both these approaches to usual CM or to each other, 2) investigated them in alcohol users or adults with SMI, 3) assessed their relative economic viability, or 4) identified which patients are more or less likely to respond to each approach.

Therefore, we will compare the efficacy of these 2 adaptations to that of the CM intervention (Usual CM) employed in the initial funding period, as well as to one another, in a sample of **240 heavy drinkers** with SMI recruited from 2 large treatment agencies. Participants with mean EtG > 499 ng/mL during a 4-week observation period will receive treatment as usual and be randomized to either a) 4 months of CM with standard reinforcement levels for submitting alcohol-negative samples (EtG < 100 ng/mL) (**Usual CM**); b) 4 months of CM with high-magnitude reinforcement for alcohol-negative samples (**High-Magnitude CM**); or c) 1 month of standard-magnitude CM for urine samples indicating light drinking (EtG < 500 ng/mL) followed by 3 months of standard-magnitude CM for alcohol-negative samples (**Shaping CM**). The primary outcome will be EtG-verified alcohol abstinence during the final 3 months of treatment (when all reinforcement is contingent on abstinence) and the 6 months of follow-up. We will also assess the impact of CM adaptations on secondary outcomes.

Because the **primary barrier to CM implementation is related to cost**,[7, 8] we will conduct a cost-benefit and cost-effectiveness analysis to assess the relative economic viability of each CM intervention. **[**Finally, little is known about moderators of outcomes in CM because past studies focused on moderators in isolation, without a theoretical model. We will use the NIAAA Addictions Neuroclinical Assessment (ANA)[33] framework to determine how its 3 domains – **executive functioning, negative emotionality, and alcohol-related incentive salience** – **predict treatment outcomes across CM adaptations.]** This study promises to have a high impact because it will determine if CM adaptations can improve outcomes and reduce costs, and identify moderators of CM efficacy in a high-cost, complex, and understudied population. Our **Specific Aims** are to:

1. Determine whether levels of alcohol abstinence during the last 3 months of treatment and a 6-month follow-up period vary by CM condition. *We hypothesize that participants in the 2 modified CM conditions will attain higher rates of abstinence than those in the Usual CM condition.*
2. Determine whether randomization groups differ on secondary alcohol outcomes, drug use, psychiatric severity, HIV risk behavior, and cigarette smoking. *We hypothesize that participants in the 2 modified CM conditions will attain better outcomes than those in the Usual CM condition.*
3. Conduct cost-benefit and cost-effectiveness analyses to determine the relative economic viability of the 3 CM conditions. *We hypothesize that the 2 modified CM conditions will dominate the Usual CM group, and the Shaping CM condition will be the most cost-effective.*
4. **[**Identify ANA-based moderators of CM treatment response across the sample and within CM conditions. *We hypothesize that higher executive dysfunction, negative emotionality, and alcohol-related incentive salience will be associated with lower levels of alcohol abstinence across all CM conditions. We also hypothesize that participants with higher levels of these characteristics will be more likely to respond to the modified CM conditions than to the Usual CM condition.***]**

2. BACKGROUND, SIGNIFICANCE, AND NEED

A. Improving Treatment for Alcohol Use and Serious Mental Illness

2.A.1 Alcohol Use Disorders (AUDs) are Common and Lead to Poor Outcomes. Up to 46% of people with SMI, defined as schizophrenia spectrum, bipolar, and recurrent major depressive disorders, suffer from an AUD in their lifetimes, **[**with 13% (schizophrenia) to 26% (bipolar disorder) suffering from an AUD in the last year.[10, 34-36]**]** Relative to people with SMI who are not substance users, those who use alcohol and drugs experience higher levels of psychotic symptoms, use of inpatient psychiatric and emergency care, medical expenditures, homelessness, medication noncompliance, treatment attrition, suicidal behavior, cognitive impairment, and violence.[11-16, 37-41]

2.A.2 Few Adults with SMI and AUDs Receive Effective Treatment. Few people receive treatment for co-occurring AUDs and SMI, and still fewer receive evidenced-based treatments.[42, 43] Cognitive behavioral therapy, motivational interviewing,[44-46] intensive case management,[47] or a combination of these treatments[48] are associated with reductions in alcohol and drug use in adults with SMI, with effects comparable to non-SMI populations.[17] However, these treatments are not widely available because of their relatively high cost, organizational barriers, and the need for extensive training of supervision of providers.[49-51] Less costly and more feasible interventions are needed to improve outcomes.

2.A.3 Contingency Management (CM) is Effective for Adults with SMI. CM interventions provide tangible reinforcers (i.e., vouchers, prizes) when patients demonstrate abstinence from substance use, typically with a urine test.[52] In 5 meta-analyses, CM was associated with increased abstinence from illicit drugs and nicotine.[52-56] Indeed, CM demonstrated larger reductions in drug use than other psychosocial treatments.[54] Although no economic analyses have been conducted on CM for AUDs, prior economic analyses of CM for illicit drug and cigarette smoking, indicate that CM is likely a good value.[1, 57-60] In one economic study usual CM ($240 in reinforcers) was found to be more cost-effective than a low-magnitude CM intervention ($80 in reinforcers), suggesting an association between a higher reinforcer magnitude and cost-effectiveness.[60]

[Several studies have shown that CM is highly acceptable to addiction providers, and our own research indicates that 75% of community mental health clinicians would use CM if it was available.[7, 61] CM is increasingly being implemented in clinical practice. For example, it is now available in the Veterans Affairs system.[62] At present, the chief barrier to CM implementation is the cost of CM reinforcers and urine tests.[7, 8]**]**

CM was associated with increased drug abstinence in 3 randomized controlled trials (RCTs) conducted in SMI populations.[19, 48, 63] One RCT (N=129) used a cognitive behavioral intervention, including CM, to treat outpatients with substance use disorders and SMI. Those assigned to the treatment condition had higher rates of drug abstinence, better quality of life, and fewer inpatient treatment episodes than controls.[48] In a study of 41 adults with SMI and substance use disorders, we found that those who received access to Social Security benefits that were contingent on alcohol and drug abstinence and related behaviors achieved higher rates of abstinence than controls.[63] In an RCT of 176 SMI outpatients receiving CM for stimulant drug abstinence, we observed that CM participants were 2.4 times more likely than controls to submit a stimulant-negative urine sample during treatment.[19] They also had lower levels of alcohol use, injection drug use, cigarette smoking, and psychiatric symptoms, and they were **5 times less likely to experience psychiatric hospitalization**.[19, 64] Differences in stimulant use continued during a 3-month follow-up period. Notably, these benefits were obtained without additional cost to providers or payers.[58] Furthermore, participants who submitted a **stimulant-positive sample before randomization attained a shorter duration of abstinence during CM** than participants whose pre-randomization samples were negative.[65]

Few studies have assessed the efficacy of CM for AUDs, given the absence of a suitable biomarker to confirm alcohol abstinence.[66] In our initial funding period we overcame this barrier by using EtG to verify abstinence. Results of the RCT for AUDs during our initial funding period provide further evidence of the efficacy of CM for adults with SMI (see Section 2.B.7).

2.A.4 Usual CM is Ineffective for Severe Addictions. Usual CM interventions provide reinforcers of relatively low value for each urine sample indicating abstinence, with a total reinforcer value of $250-$400 over 12 weeks.[30] Substance use immediately before CM, as verified by urine drug test, is a consistent predictor of poor response to usual CM.[65, 67-69] In a RCT of usual CM for cocaine dependence, participants who submitted a cocaine-positive urine sample before CM attained a median of $44 in CM reinforcers, which was 5 times less than participants with negative pre-treatment samples (median= $237).[67] In another study, those who tested positive for stimulants at study entry submitted stimulant-positive urine samples 64% of the time during usual CM, while those who were stimulant-negative submitted positive samples only 18% of the time.[68] Our study described in Section 2.A.3[65] and our preliminary data from the initial funding period (Section 2.B.7) also found an association of pre-treatment urine tests with outcomes in usual CM.

2.A.5 CM Must be Modified to Treat Heavy Drinkers with SMI. Initial studies of CM in cigarette smokers and cocaine users showed that drug abstinence can be more effectively attained either by increasing the magnitude of reinforcement (high-magnitude) or by reinforcing successively closer approximations of

abstinence (shaping), instead of requiring abstinence only.[26-32, 69-71] In previous studies, **high-magnitude CM** (i.e., higher value of reinforcers) improved addiction treatment outcomes, particularly for those with recent pre-treatment substance use.[30-32, 71] While not all studies have found a link between high-magnitude CM and increased abstinence,[72, 73] null results might stem from the use of samples that included participants with low-severity addictions, who might not benefit from high-magnitude CM.

Four studies investigated the efficacy of high-magnitude CM in people with severe addictions. In one, 29 non-responders to a CM intervention for cocaine were exposed to high-magnitude CM (up to $3,480) and usual CM (up to $382). During high-magnitude CM, 45% attained 4 or more weeks of abstinence, while only 5% achieved this goal during standard-magnitude CM.[32] In another study, high-magnitude CM increased drug abstinence in 11 treatment-resistant cocaine and opioid users.[31] In both of these studies, participants also submitted more opiate- and benzodiazepine-negative urine samples during high-magnitude CM relative to usual CM or usual care conditions. In a third study, 109 participants who submitted a pre-treatment cocaine-positive urine sample were randomized to usual CM (reinforcer value $240), or high-magnitude CM (reinforcer value $560). For those assigned to high-magnitude CM, the duration of abstinence was 2.5 times longer than for those in usual care and 1.6 weeks longer than for those in standard CM. However, the latter difference was not statistically significant.[30] In a fourth study, conducted by our group, we found that among 103 cigarette smokers, high-magnitude CM and pre-CM smoking-negative saliva tests were both correlated with higher rates of smoking abstinence during CM.[69, 71]

Shaping CM involves reinforcing reductions in use before requiring abstinence. It is associated with better outcomes in people who do not respond to abstinence-based CM.[26-28] One study randomized 95 adults to receive either 8 weeks of CM for cocaine abstinence or 3 weeks of a shaping CM condition, in which they were initially required to reduce cocaine metabolite levels by 25% to receive reinforcers, and then received 5 weeks of CM for total abstinence.[27] After 8 weeks, participants in the shaping condition had higher rates of abstinence than the abstinence-only group. A series of studies, conducted by Lamb and colleagues supports the efficacy of shaping interventions for hard-to-treat cigarette smokers.[26, 28, 29] In these studies, pre-intervention carbon monoxide readings were used to personalize targets for CM in shaping schedules. In an RCT of 96 hard-to-treat smokers, participants who received shaping CM based on initial carbon monoxide results submitted 6 times as many smoking-abstinent samples as those who received abstinence-based CM.[29]

2.A.6 NIAAA's Addictions Neuroclinical Assessment (ANA) Framework. **[**While CM is among the most effective interventions for substance use disorders, little information is available to predict treatment response. As noted in Sections 2.A.4 and 2.B.7, pre-treatment substance use predicts poor CM response. Other predictors of poor treatment response include withdrawal intolerance, a history of legal involvement, and low levels of motivation, as well as high levels of stressful events, passive coping, psychiatric symptoms, impulsivity, and non-adherence to other treatments.[65, 74-78] However, previous studies have **investigated these variables in isolation** without a theoretical framework.

Recently, NIAAA leadership proposed the ANA, a neuroscience-based framework for addiction.[33] This approach postulates 3 domains – poor executive functioning (e.g., working memory, impulsivity), negative emotionality (e.g., depression, anxiety, psychological symptoms of withdrawal), and high levels of alcohol-related incentive salience (e.g., thinking about alcohol, craving a drink) – as the primary factors that cause and maintain AUDs. The authors propose that ANA domains, assessed by self-report and cognitive testing, as well as genetic analysis, and eventually neuroimaging can be used to identify groups of patients who do not respond to existing treatments. The ANA also provides a framework to guide treatment adaptations to improve outcomes for specific groups of non-responders (e.g. those with high levels of executive dysfunction). It might also be used to match individuals to specific interventions based on their ANA characteristics. Therefore, the ANA provides a structure for developing a personalized medicine approach to AUD treatment.

The ANA is particularly relevant to CM in our priority population, because 1) heavy drinkers and adults with SMI have high rates of executive dysfunction[79] and negative emotionality; 2) heavy drinkers have high levels of alcohol-related incentive salience;[80] 3) combating alcohol-related incentive salience with an alternate reinforcer is the basis of CM;[81] 4) little research has described the impact of cognitive dysfunction on treatment outcomes for co-occurring disorders;[17] and 5) we have observed that negative mood states are associated with poor outcomes in CM among adults with SMI, particularly in heavy drinkers (Section 2.B.8).[65] Our overall hypothesis is that heavy drinkers with higher levels of executive dysfunction, negative emotionality, and alcohol-related incentive salience are more likely to respond to modified CM conditions than to Usual CM. We also hypothesize that ANA domains will be better predictors of treatment response in specific CM adaptations. For example, those with high alcohol-related incentive salience might respond better to the proposed High-Magnitude CM than to Shaping or Usual CM conditions, because a highmagnitude of reinforcers will likely be necessary to overcome high levels of alcohol-related incentive salience (i.e., reinforcing effects of alcohol).**]**

B. Progress Report and Preliminary Studies

2.B.1 Beginning & Ending Dates. Funding period: 3/1/12 to 2/28/15, no-cost extension through 7/31/16.

2.B.2 Specific Aims. *Aim 1:* To conduct an RCT to determine whether a 12-week CM intervention that provides reinforcement for alcohol abstinence (measured by EtG) and treatment attendance would be successful in reducing alcohol use in 79 people with alcohol dependence and SMI who received treatment as usual. *Aim 2:* To determine whether the CM intervention succeeded in improving addiction treatment attendance. *Aim 3:* To determine whether the CM intervention led to improvements in related outcomes.

2.B.3 Significance. This was the first adequately powered RCT of CM that used EtG, an accurate and feasible alcohol biomarker, to treat AUDs, and only the second RCT of CM in people with SMI.

2.B.4 Participant Enrollment and Retention. A total of 121 adults met inclusion criteria. Eighty-four (69%) completed the 4-week induction period and 79 were randomized to CM (n=40) or Non-Contingent (NC) control (n=39) conditions. Twenty-six CM participants (65%) and 29 NC participants (74%) completed treatment. Thirty CM participants (75%) and 30 NC participants (77%) completed at least one follow-up interview, representing an improvement in retention relative to our previous SMI study, particularly in the CM group. Better retention was likely due to our 4-week induction period.

2.B.5 EtG and Self-Reported Alcohol Use. We found that an EtG > 499 ng/mL detects 71% of heavy drinking in the past 2 days and 58% of light drinking in the same period, but only 59% of heavy drinking in the past 5 days. In contrast, EtG > 100 ng/mL detects 77% of light drinking in the past 2 days and 79% of heavy drinking in the past 5 days.[21] Therefore, an **EtG level > 499 ng/mL equates to recent heavy drinking**, while a result of **100-499 ng/mL equates to recent light drinking** or heavy drinking during the last 3-5 days.

2.B.6 CM Is Effective. As described in the Appendix, CM participants were **3.1 times** (95% CI=2.2-4.5) more likely than controls to submit **alcohol-negative samples** during treatment.[25] **The duration of alcohol abstinence** (defined as the number of consecutive alcohol-negative EtG samples) **was twice as long in the CM group as in controls** (CM M=8.6, SD=12.6, controls M=4.1, SD=1.2, $p<0.05$). The mean EtG level (409 ng/mL, SE=40 ng/mL) in CM participants was lower than in controls (735 ng/mL, SE=41 ng/mL, $p< 0.05$). CM participants reported fewer days of alcohol use in treatment (M=3.7, SE=1.6) and follow-up (M=3.3, SE=1.4), compared to controls (treatment M=12.0, SE=1.5, $p<0.05$, follow-up M=10.7, SE=1.4, $p<0.05$). Those in CM were less likely to than controls to report heavy drinking in treatment (OR=3.5, 95% CI=2.3-5.2) and follow-up (OR=4.9, 95% CI=1.8-13.5). CM participants were more likely than controls to submit stimulant-negative urine samples during treatment (OR=3.2, 95% CI=2.0-5.1) and follow-up (OR=4.6, 95% CI=1.2-17.1), and smoking-negative carbon monoxide breath samples during treatment (OR=5.4, 95% CI=1.9-15.0).

2.B.7 Heavy Drinkers do Not Respond to Usual CM. The efficacy of CM was associated with average EtG levels during a 4-week induction (i.e., pre-randomization) period. CM participants with mean pre-randomization EtG < 500 ng/mL (light drinking) attained continuous abstinence that was **2.3 times longer** than those with mean EtG > 499 ng/mL (recent heavy drinking; p=0.03; see Table 1). This equals 2 weeks of additional abstinence. Light drinkers also attained a duration of continuous drug abstinence that was 5 times longer (M=10.7, SD=9.5) than heavy drinkers (M=1.9, SD=3.0). Finally, CM participants with pre-treatment EtG > 499 ng/mL had levels of abstinence similar to controls, indicating that they did not respond to CM.

2.B.8 Depression is Linked to Poor Outcomes in Heavy Drinkers. **[**As predicted by the ANA, heavy drinkers (n=42, EtG > 499 ng/mL) with a major depressive disorder attained an average of 1 EtG-negative urine sample during treatment (SE=2.2), whereas non-depressed participants averaged 5.6 consecutive EtG-negative samples (SE=1.5; $p<0.05$). Among heavy drinkers in CM (n=16), those with depression attained less abstinence than non-depressed participants. The difference was not statistically significant, likely due to insufficient power. Notably, **depression was not associated with alcohol abstinence in light drinkers**. These findings support our premise that ANA domains can predict alcohol abstinence in heavy drinkers.**]**

C. Summary of Significance and Scientific Premise

[Our long-term goal is to create a personalized medicine approach to CM in which patients are matched to a specific type of CM intervention based on their pre-treatment clinical characteristics (e.g., EtG level, ANA variables). CM is an effective intervention, and psychosocial treatments have been shown to improve outcomes for patients with co-occurring disorders.[17, 54] However, many patients do not respond to these treatments. The absence of rigorous research on how to improve the efficacy of treatments for co-occurring disorders is a major barrier to better care for this population.[17, 18] Therefore, by investigating the efficacy of CM adaptations, we will address a major gap in the literature on CM and on the treatment of co-occurring disorders. The significance of our approach is enhanced by the fact that heavy drinkers with SMI – a **high cost, complex, and understudied population** – have **very limited options for effective outpatient care**. As a result, they rely on costly acute and inpatient care.[13] If results support our hypotheses, modified CM approaches could offer **heavy drinkers with and without SMI** an effective treatment option.

The premise for **Aim 1** is robust because 1) our pilot data demonstrate that those who engaged in heavy drinking before randomization did not respond to usual CM, 2) previous studies support the efficacy of high-magnitude and shaping CM in illicit drug users and smokers, 3) these studies had methodological weaknesses (e.g., small samples sizes), 4) they did not focus on AUDs, and 5) they did not compare high-magnitude and

shaping CM. Determining the efficacy of shaping CM might have particular relevance for AUDs, relative to illegal drugs or cigarette smoking, because reductions in drinking (especially heavy drinking days) lead to clinical and cost benefits.[82, 83] The premise for **Aim 2** is based on our studies that demonstrate secondary effects of CM on non-targeted drug use, cigarette smoking, psychiatric symptoms, and hospitalization. The premise for **Aim 3** is based on the limited economic literature suggesting that CM is a "valuable" approach, and one study that observed a relationship between reinforcer magnitude and cost-effectiveness.[60] No economic studies have addressed AUDs, or treatment non-responders, or included a comprehensive, meaningful set of measures of outcomes and resource utilization, or compared proposed adaptations to Usual CM. We propose comprehensive measures of study and non-study resources. Our "real-world" price weights will offer stakeholders an accurate, meaningful estimate of the "true costs" of CM adaptations for a high-need group. We propose that the Shaping CM will dominate the other conditions, because shaping may lead to increased clinical benefit without increasing the costs of reinforcers. The economic analysis will affect both policy and practice, since the primary barrier to CM implementation is the cost of reinforcers and urine tests.[7, 8] The premise for **Aim 4** is described in detail in Section 2.A.6; in brief, the ANA provides a framework for identifying subgroups of heavy drinkers who are more or less likely to respond to the proposed CM interventions.**]**

3. INNOVATION

[Our proposal is highly innovative. **First**, it will be one of very few CM studies to investigate how CM can be modified to improve outcomes for non-responders, and the only one to focus on alcohol. **Second**, it will be the first study to determine whether modifications to a co-occurring disorders treatment improve outcomes in non-responders. **Third**, it will be the first adequately powered RCT of a high-magnitude CM intervention to treat non-responders. Previous research involved within-subject comparisons or insufficient sample sizes to detect differences between low- and high-magnitude approaches.[30-32, 71] **Fourth**, it will be only the second RCT to compare shaping CM to abstinence-based CM interventions; notably, the first addressed tobacco instead of alcohol. Our focus on a shaping CM intervention for AUDs is therefore highly innovative, and it might have significant clinical applications for patients who are initially unwilling to become abstinent. **Fifth**, it will be the first RCT to compare outcomes between high-magnitude and shaping CM conditions. **Sixth**, we propose CM adaptations that are clinically feasible, rather than approaches such as percentile schedules that might be overly burdensome for clinicians. **Seventh**, our economic analysis is novel, as it will be the first comprehensive economic evaluation of CM for AUDs. In addition, we will include cost-benefit and cost-effectiveness analyses of our CM adaptations for multiple stakeholders (payers, providers, society) and multiple measures of effectiveness: quality-adjusted life-years (QALYs), time abstinent, and heavy drinking days averted (see Section 4.G.2). QALYs are an important economic measure, yet only 2 extant CM economic evaluations have incorporated them as an outcome measure, and we conducted both of them.[1, 58] **Eighth**, the addition of heavy drinking days averted as an economic outcome is novel and timely, as this outcome is highly valued by clinical stakeholders and policymakers. **Ninth**, this will be the first CM RCT to investigate a set of treatment modifiers using a theoretical framework (ANA) relevant to heavy drinkers, adults with SMI, and the CM adaptations we propose. While ANA variables have been investigated before, **ours will be the first alcohol treatment study to investigate the ANA as a framework for predicting outcomes**. If results support our hypotheses, clinicians will be able to use **pre-treatment EtG levels and ANA variables to match patients with the type of CM that is most likely to achieve positive outcomes**. In addition, the proposed study will be the first multi-site CM trial of adults with SMI, one of few CM studies to use a follow-up longer than 3 months, and the first to use the NIH Toolbox in adults with co-occurring disorders.**]**

4. APPROACH

A. Investigative Team

Michael McDonell, PhD (PI), is an Associate Professor in the Elson S. Floyd College of Medicine and the Associate Director of the Initiative for Research and Education to Advance Community Health at Washington State University (WSU). He is the Principal Investigator (PI) or Project Lead of 3 RCTs of CM as a treatment for AUDs. His research uses the EtG biomarker to implement and evaluate CM in adults with SMI, as well as in American Indians and Alaska Natives. His current research involves 5 sites across the US. He has authored 26 peer-reviewed articles and obtained 5 externally funded grants in collaboration with the Co-Investigators on this application.

Table 2. Expertise of study investigators

Expertise	McDonell	McPherson	Murphy	Roll	Chaytor	Ries
CM for Alcohol	**Lead**	√		√		√
Ethyl glucuronide	**Lead**	√				√
Statistics/Methods		**Lead**	√	√		
Cost Analysis			**Lead**			
CM Schedules	√	√		**Lead**		
Assessment					**Lead**	
Co-Occurring SMI	√			√		**Lead**

Sterling McPherson, PhD, is an Assistant Professor and the Director of Biostatistics in the College of Medicine at WSU and Associate Director for Analytics at Providence Medical Center. He has authored more than 55 publications and was the lead analyst for 9 CM studies, including those led by Dr. McDonell. Dr. McPherson investigates behavioral and pharmacological treatments for addictions, missing data analyses, and innovative longitudinal methods. He will assist in protocol development and oversee all data analyses.

Sean Murphy, PhD, is an Associate Professor in the Department of Health Policy and Administration at WSU. He is a health economist specializing in economic analyses in substance use treatment trials and has conducted economic analyses of CM.[1, 58] Dr. Murphy will design and conduct all economic analyses.

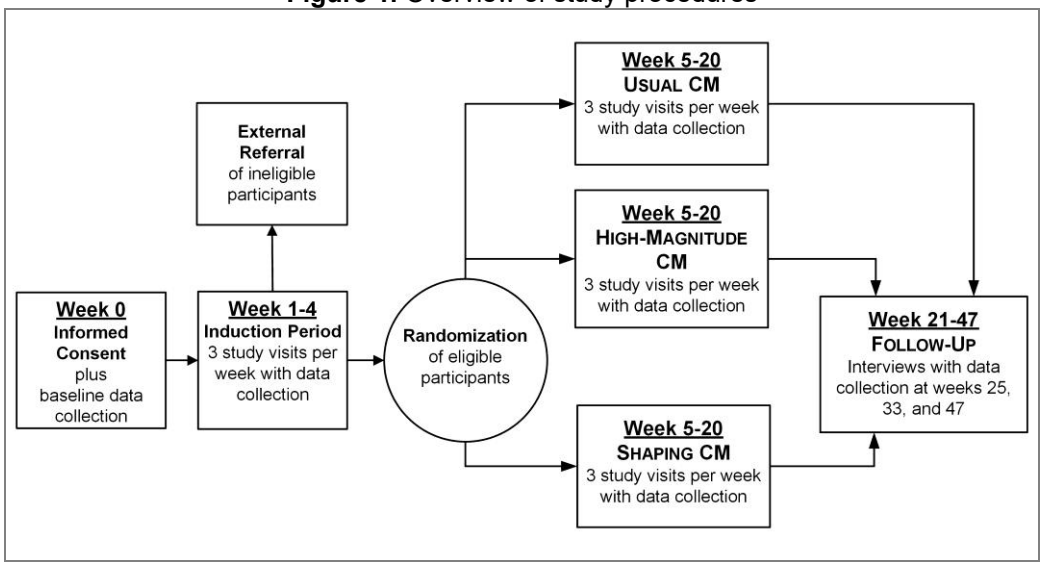
John Roll, PhD, is a Professor and Vice Dean of Research in the College of Medicine at WSU. He has been a PI of an NIH Center of Excellence and a PI or Co-Investigator on many large CM studies. He will provide expertise in developing CM schedules of reinforcement to maximize reductions in substance use.

Naomi Chaytor, PhD, ABPP, is an Associate Professor in the College of Medicine at WSU. She is a board-certified clinical neuropsychologist whose research focuses on cognitive and psychological predictors of treatment outcomes in adults with chronic conditions. She will oversee training, quality control, and interpretation of all cognitive and emotional assessment instruments.

Richard Ries, MD, is a Professor of Psychiatry at the University of Washington with more than 30 years of experience conducting research on co-occurring disorders, including serving as a PI on 2 RCTs of CM in SMI populations. Dr. Ries will provide extensive expertise in treating co-occurring disorders, serve as the study medical director, supervise the Seattle site, and assist with dissemination of results.

B. Overview of Study Activities

Figure 1. Overview of study procedures



We propose a **highly rigorous** evaluation of 3 CM interventions for heavy drinking in a sample of 240 adults with SMI. To recruit a sufficient number of recent frequent heavy drinkers (mean EtG > 499 ng/mL), we anticipate the need to recruit 400 participants altogether. All will receive psychiatric and addiction treatment as usual. As shown in **Figure 1**, these 400 participants will first take part in a 4-week induction period during which they will receive reinforcers for submitting 3 urine samples per week, regardless of EtG results.

Participants who meet secondary eligibility criteria (N=240; see Section 4.D.2) will be randomized to receive either a) 4 months of standard-magnitude reinforcement CM for submitting alcohol-negative samples (EtG < 100 ng/mL) (**Usual CM**); b) 4 months of high-magnitude CM for submitting alcohol-negative samples (**High-Magnitude CM**); or c) 1 month of standard-magnitude CM for submitting urine samples that indicate light drinking (EtG < 500 ng/mL), followed by 3 months of standard-magnitude CM for submitting alcohol-negative samples (**Shaping CM**). Our CM paradigm will use the variable magnitude of reinforcement procedure (VMRP; see Section 4.E.2), in which participants draw from a bowl for chances to receive items and gift cards. Groups will differ only on the number of draws they receive (Usual vs. High-Magnitude), or the contingency by which they are allowed to engage in draws (light drinking vs. abstinence).

Randomized participants will complete follow-up assessments at 1, 3, and 6 months to assess long-term outcomes. The primary outcome will be alcohol abstinence, assessed as EtG < 100 ng/mL, during the last 3 months of treatment (when all reinforcers are contingent on abstinence) and the 6-month follow-up period. We will describe group differences in heavy drinking, other alcohol outcomes, drug use, psychiatric symptoms, and health behaviors. We will also estimate the costs associated with implementing and sustaining each CM intervention in a clinical setting, and determine which condition provides the most economic value. Finally, we will assess whether ANA variables are useful in predicting alcohol abstinence during the 3 CM interventions.

C. Sampling and Recruitment Plan

4.C.1 Sampling Plan. **[**The sample will be drawn from 12,000 outpatients with SMI receiving care at 2 large mental health organizations: Community Psychiatric Clinic in Seattle, WA (metro population ~3 million) and Frontier Behavioral Health in Spokane, WA (metro population ~666,000; see Letters of Support). Each site

reports that about 25% of their consumers (3,000 overall) suffer from AUDs. Based on screening conducted in our previous study, we estimate that 60% of this subsample (1,800) will meet criteria for heavy drinking during Year 1 of the proposed study. The two study sites anticipate a total of 500 new patients who are heavy drinkers each year. This equates to an overall recruitment pool of 3,050 over 3.5 years (1,800 in Year 1 and then 500 new consumers annually for 2.5 years). We will recruit 13% (N=400) and randomize 8% (n=240) of eligible heavy drinkers to meet our goals. In our initial funding period, 1.5 research staff recruited 121 participants at 1 site over 2 years, and 84 met criteria for randomization. This equates to 40 consented and 28 randomized participants per year per full-time staff effort. For this proposal, 3.65 research staff (more than twice as many) will recruit participants at 2 sites over 3.5 years. Using these parameters, we will be able to recruit 515 participants and randomize 357. These numbers exceed our recruitment goals by 23% and 33%, respectively. If more than 400 heavy drinkers are needed to randomize 240, we will continue recruitment until we reach our goal. If accrual is much lower than anticipated, Drs. McDonell and Ries will approach other community mental health centers where they collaborate on other studies and arrange to expand recruitment to those sites.**]**

4.C.2 Recruitment, Screening, and Informed Consent. Our recruitment and consent procedures were developed and refined during our previous RCTs of CM with this population **[**and our community based-RCTs with other populations.**]** Recruitment materials will be placed in clinic waiting rooms, drop-in centers, and other locations where potential participants congregate. Clinic staff will distribute study recruitment materials that ask consumers if they are interested in participating in the study. Research staff will also offer regular recruitment presentations at addictions and mental health treatment groups. **[**We will also expand recruitment to other agencies, shelters, detox facilities, and hospitals that serve patients with co-occurring disorders; post flyers throughout Seattle and Spokane; and advertise in newspapers (including newspapers for homeless adults), Craigslist, Facebook, websites, radio, and press releases.**]**

Interested patients who contact study staff will be screened. Potentially eligible patients will be scheduled for an in-person baseline appointment, during which study staff will explain the purpose of the study and the procedures involved. Next, they will administer a brief screening measure that assesses capacity to consent to research procedures.[84] If results indicate concerns about competency to provide consent, the potential participant will complete the MacArthur Competency Assessment Tool for Clinical Research (MacCAT-CR), the gold standard for assessing informed consent capacity.[85] Only those who demonstrate capacity to provide informed consent will be enrolled. Participants will review and sign the consent form in the presence of study staff, who will read the form aloud as needed.

D. Eligibility and Randomization

4.D.1 Initial Eligibility. Eligibility criteria were selected to enable accurate assessment of intervention efficacy while maximizing the generalizability of study results. The PI will make final decisions regarding eligibility. **Inclusion criteria** are 1) 4 or more standard drinks on 5 or more occasions in the prior 30 days; 2) DSM-5 [86] diagnosis of AUD, moderate to severe; 3) DSM-5 diagnosis of schizophrenia or schizoaffective disorder, bipolar I or II, or recurrent major depressive disorder (>1 episode); 4) age 18-65 years; and 5) receipt of, or eligibility to receive, treatment as usual at study sites. **Exclusion criteria** are 1) current DSM-5 diagnosis of substance use disorder, severe; 2) inability to demonstrate competency to provide consent on MacCAT-CR; 3) risk of medically dangerous alcohol withdrawal (i.e., seizure within the last 12 months, concern by participant or clinician regarding a potentially dangerous withdrawal); 4) prior diagnosis of dementia; and 5) determination by the PI and Dr. Ries (medical director) that participation would be medically or psychiatrically unsafe.

4.D.2 Induction Period. Eligible participants will take part in a 4-week induction period, during which they will engage in the VMRP procedure 3 times a week (Section 4.E.2) for providing urine samples. At each visit, they will receive 3 draws for prizes when they provide urine samples, regardless of whether the samples are positive for alcohol use. Those who provide at least 1 urine sample during each of these 4 weeks will receive a $20 bonus incentive. Consistent with previous studies using this approach,[87] only participants who demonstrate that they are frequent heavy drinkers and **[**continue to attend study visits**]** will be randomized. In our initial funding period, use of the induction period resulted in a CM completion rate of 65%, which exceeds published CM retention rates in non-SMI populations. Therefore, participants who 1) **attain an average EtG score of > 499 ng/mL** (indicating recent heavy drinking) and 2) **attend at least 1 study visit during the final week of the induction period** will be randomized. Given pilot data indicating that 53% of our sample during the initial funding period (n=42) had a mean induction EtG > 499 ng/mL, as well as our screening procedures to rule out light drinkers, we anticipate that at least 60% of the 400 consented participants (n=240) will be eligible for randomization. We will continue to recruit participants until 240 participants are randomized. Participants who do not meet criteria for randomization will be referred to other available AUD treatments.

4.D.3 Randomization Procedures. We will randomize 240 participants to treatment conditions by using permuted block randomization. We will stratify across the following variables: 1) study site, 2) gender, and 3) baseline EtG level > 1,000 ng/mL, which indicates very heavy recent drinking. Dr. McPherson will randomize all participants to treatment conditions and will not have direct contact with participants.

E. Study Interventions

4.E.1 Treatment as Usual. The **Community Psychiatric Clinic** (CPC) in Seattle is a dually licensed chemical dependency and mental health agency with 4 large clinics, as well as other locations throughout Seattle. Its mental health program provides case management, medication management, group and individual counseling, vocational services, and housing services based on the needs of each consumer. Consumers can be seen daily for services, but are more often seen once or twice per month. CPC provides intensive outpatient addiction treatment, with a focus on treating adults with co-occurring mental illness. Addiction groups are 2 to 3 times per week. Integrated dual disorders treatment is also available. Addiction providers use cognitive behavioral therapy, motivational interviewing, and 12-step facilitation. **Frontier Behavioral Health** (FBH) is the largest provider of mental healthcare in Spokane County. At its 8 locations, FBH provides case management (including assertive community treatment), medication management, group psychotherapy, crisis management, and supported employment. Consumers can be seen as often as daily, but typically see a clinician once or twice per month, depending on need. While FBH is not a certified chemical dependency provider, it offers an integrated dual disorders program that provides group and individual therapy. Clinicians use cognitive behavioral therapy, motivational interviewing, and 12-step facilitation. Consumers who suffer from substance use disorders but are not enrolled in this program are referred to local addiction agencies.

4.E.2 Variable Magnitude of Reinforcement Procedure (VMRP). Participants in the 3 CM conditions will engage in VMRP each time they meet criteria for obtaining reinforcers over the 16-week treatment period. Groups will differ only by the number of times they engage in VMRP (Usual CM vs. High-Magnitude CM) or the criterion required to receive reinforcement (light drinking vs. abstinence). VMRP will involve drawing from a bowl of 500 chips, some of which represent prizes with different dollar amounts. Fifty percent will say "good job" (no prize), 41.8% will result in a small prize ($1 to $5 value), 8% will result in a large prize ($20 to $30 value), and 0.2% will result in a jumbo prize ($80 to $100 value). For each urine sample that meets reinforcement criteria, participants will complete the corresponding number of draws. **Table 3** summarizes the available number of prize draws and earnings, both maximum and expected, for each CM condition. The actual maximum payout will depend on urine test results and the percentage of study visits attended. The number of drawings will escalate with each full week (3 consecutive samples) that meets criteria for reinforcement. Missing samples or samples that do not meet reinforcement criteria will result in a reset to the original number of draws (3 draws for Usual and Shaping CM, 14 draws for High-Magnitude CM) on the next occasion when the participant provides a sample that meets reinforcement criteria. If a reset occurs, participants must provide 3 consecutive samples that meet reinforcement criteria before they are returned to the number of draws they previously accumulated. This procedure reduces the probability of relapse after abstinence is initiated.[88]

4.E.3 Types of Reinforcers. Prizes will be displayed in cabinets in study offices at clinic sites. They will be selected on the basis of participant feedback in our previous studies with this population, and will likely include toiletries, small clothing items, gift cards, mp3 players, DVD players, digital cameras, and smartphones.

4.E.4 Usual CM Condition. Participants in the Usual CM condition will receive the same CM intervention administered in the initial funding period and in numerous other studies.[55] They will earn at least 3 VMRP draws each time they submit an alcohol-negative urine sample (EtG < 100 ng/mL), and they will receive an additional draw for each week (3 consecutive alcohol-negative samples) of abstinence. Continuously abstinent participants will receive 18 draws for alcohol-negative tests at each of their week 20 appointments. Missing or alcohol-positive samples will result in a reset to 3 draws, as explained above in Section 4.E.2.

4.E.5 High-Magnitude CM. Participants in this group will be required to demonstrate alcohol abstinence (EtG <100 ng/mL) to receive reinforcers. However, they will receive twice as many VMRP prize draws overall relative to the Usual CM group for submitting alcohol-negative samples during the 16-week treatment period, since they will earn at least 14 draws for each negative sample. Note that, although the initial number of draws in the High-Magnitude group is about 3 times higher than in the Usual CM group, the difference in the number of draws will become less pronounced over the 16-week treatment period, and will ultimately approximate twice as many chances to engage in VMRP (see Table 3). One additional draw will be accumulated for each week (3 continuous alcohol-negative samples) of abstinence. Continuously abstinent participants will be able to earn 29 draws for alcohol-negative samples at each appointment in week 20. Missing or alcohol-positive samples will result in a reset to 14 draws, as explained above in Section 4.E.2.

Table 3. Summary of CM schedules and maximum earnings for randomized groups

CM Group	*Criteria for Reinforcement*	*Maximum # of Prize Draws*	*Maximum/ Expected Payout*
Usual	EtG< 100 ng/mL (abstinence) in weeks 5-20	504	$1,097/$500
High-Magnitude	EtG< 100 ng/mL (abstinence) in weeks 5-20	1032	$2,248/$1,000
Shaping	EtG< 500 ng/mL (light drinking) in weeks 5-8 EtG< 100 ng/mL (abstinence) in weeks 9-20	504	$1,097/$500

4.E.6 Shaping CM. Participants will receive the same number of VMRP prize draws as the Usual CM group. However, during the first 4 weeks of treatment (weeks 5-8), the criterion for engaging in VMRP will be light drinking (EtG < 500 ng/mL) rather than abstinence. During the final 3 months of treatment (weeks 9-20) the criterion for receiving reinforcers will shift to alcohol abstinence (EtG < 100 ng/mL). Participants will earn at least 3 draws for each urine sample indicating no heavy drinking during weeks 5-8 and no alcohol use at all during weeks 9-20. An additional draw will be accumulated for each consecutive week (3 consecutive samples) in which participants meet all criteria for receiving reinforcers. Those who continuously meet criteria for reinforcers will be able to earn 18 draws at each of their week 20 appointments. Missing samples or samples that do not meet criteria for reinforcement will result in a reset to 3 draws, as explained above in Section 4.E.2.

F. Data Collection and Measures

4.F.1 Schedule. Data collection will occur during the study period (weeks 1-20, including baseline assessment, induction, and treatment) and the follow-up period (weeks 21-47) as shown in **Table 4**. Baseline data collection will take ~120 minutes and include urine tests, self-reported data, and cognitive tests. Participants will provide urine samples every Monday, Wednesday, and Friday during the study period. At each visit, they will complete brief self-report questionnaires on alcohol use. Weeks 4, 8, 12, 16, and 20 will include 45-minute data collection visits to assess other outcomes. During follow-up, participants will be scheduled for 45-minute visits at weeks 25, 33, and 47. They will receive a $30 gift card for completing the baseline interview and a $20 gift card for completing each interview in weeks 4, 8, 12, 16, 20, 25, 33, and 47.

Table 4. Data collection schedule

Measure	Baseline Wk 0	Induction Wk 1-4	Treatment Wk 5-20	Follow-up Wk 25, 33, 47
Eligibility Criteria				
DSM-5 inclusion/exclusion diagnosis (interview)	√			
Consumer at participating agency (self-report)	√			
Age (self-report)	√			
Current heavy drinking (self-report)	√			
Health and risk factor exclusions (self-report/observation)	√			
Primary Outcome				
Alcohol use (ethyl glucuronide)	√	3x / Wk	3x / Wk	Every visit
Other Secondary Outcomes				
Any/Heavy drinking: ATLFB (self-report)	√	3x / Wk	3x / Wk	Every visit
Recent heavy drinking (ethyl glucuronide)	√	3x / Wk	3x / Wk	Every visit
Alcohol cravings (self-report)	√	Weekly	Weekly	Every visit
Alcohol/ drug severity: ASI-Lite (self-report)	√	Wk 4	Wk 8,12, 16, 20	Every visit
Drug use (biochemical)	√	3x / Wk	3x / Wk	Every visit
Cigarette Timeline FollowBack (self-report)	√	Wk 4	Wk 8,12, 16, 20	Every visit
Nicotine dependence: Fagerstrom (self-report)	√	Wk 4	Wk 8,12, 16, 20	Every visit
HIV risk behavior (self-report)	√	Wk 4	Wk 8,12, 16, 20	Every visit
Psychiatric symptoms: PANSS (interviewer rating)	√	Wk 4	Wk 8,12, 16, 20	Every visit
Economic Measures				
Drug Abuse Treatment Cost Analysis Program tool				Wk 47
Healthcare use by patient medical records	Past 12 mo	Throughout	Throughout	Throughout
Non-Study Resources Form (self-report)	√	Wk 4	Wk 8,12, 16, 20	Every visit
Quality-Adjusted Life-Years: EQ-5D (self-report)	√	Wk 4	Wk 8,12, 16, 20	Every visit
ANA Moderators				
Negative emotionality: NIH Toolbox (self-report)	√	Weekly		
Executive functioning: NIH Toolbox/MCQ (cognitive tests/self-report)	√			
Alcohol incentive salience (self-report/cognitive tests)	√			
Adverse Events				
Alcohol withdrawal: SHOT (interviewer rating)	√	3x / Wk	3x / Wk	Every visit
Participant safety (self-report/observation)	√	3x / Wk	3x / Wk	Every visit

4.F.2 Measures Overview. Baseline data will include assessment of eligibility criteria and measures for Aims 1-4. Our primary outcome (Aim 1) is biochemically confirmed alcohol abstinence, measured 3 times per week during the final 12 weeks of the treatment period (weeks 9-20) and 3 times total during the 6-month follow-up period. Aim 2 outcomes will be collected in a similar fashion. Cost data will be collected for 12 months before and after randomization. Aim 3 outcomes consist primarily of self-reported use of non-study

healthcare resources, QALYs, days of alcohol abstinence, and heavy drinking days averted. We will collect data on each ANA domain to test for predictors of alcohol abstinence during CM interventions.

4.F.3 Eligibility Criteria. Age and frequency of recent drinking will be assessed by phone screening. The Structured Clinical Interview for DSM-5 (SCID-5) interview, clinical trials version,[89] will be administered to assess diagnostic inclusion (SMI, AUDs) and exclusion criteria. Competency to provide informed consent will be assessed as described in Section 4.C.2.

4.F.4 Primary Outcome. Our primary outcome will be alcohol abstinence during the last 3 months of treatment and the 6-month follow-up period. At each study visit, urine samples will be collected and analyzed onsite for EtG by using the **[**commercially available Diagnostic Reagents Incorporated EtG immunoassay on an Indiko Bench Top Analyzer (ThermoFischer Scientific, Freemont, CA).**]** We will use the assay to create a binary indicator of alcohol use defined by EtG < 100 ng/mL. That threshold was used in our first funding period and in our ongoing CM trials. While it is not associated with false positives due to non-beverage alcohol exposure,[21, 22] we will ask participants to abstain from using all alcohol-containing products, like hand sanitizer.

4.F.5 Secondary Outcomes. **Alcohol use:** Days of self-reported abstinence and heavy drinking will be assessed at each visit by the Alcohol Timeline FollowBack,[90] which measures the frequency and amount of daily drinking. Recent heavy drinking will also be assessed by EtG tests with a cut-off of 500 ng/mL.[22] The Addiction Severity Index Lite[91] will assess the impact of alcohol use on psychiatric, legal, medical, and family functioning, as well as self-reported drug use and its severity. We will assess alcohol cravings by using a 10 cm visual analog scale anchored at 0 (no craving) and 100 (most intense craving possible). **Drug use:** At each study visit, urine samples will be tested for opioids, amphetamine, methamphetamine, cocaine, and cannabis with EZ-split point-of-care immunoassays. **Health-impairing behaviors:** The Timeline FollowBack method will assess the number of cigarettes smoked daily.[90] The Fagerstrom Test of Nicotine Dependence[92] will assess the presence and severity of nicotine dependence. HIV risk behavior will be assessed with the brief HIV Risk Behavior Scale.[93] **Psychiatric symptoms:** The Positive and Negative Syndrome Scale (PANSS),[94] a commonly used measure, will be used to monitor the clinician-rated severity of positive (e.g., hallucinations) and negative (e.g., avolition) symptoms of schizophrenia-type illness, as well as mood and anxiety symptoms.

4.F.6 Economic Measures. The average administrative cost of CM will be estimated by using a modified version of the Drug Abuse Treatment Cost Analysis Program (DATCAP) instrument.[95] DATCAP is a standardized, customizable tool that can estimate the costs of programs in various settings. Costs in our analysis will include labor, office and urine testing supplies, and CM reinforcers. A detailed costing method will be used to value other resources. Data on within-study resource use will be retrieved from study records. Data on use of non-study medical resources (e.g., medication and use of inpatient, outpatient, and emergency services) will be collected from patients' electronic health records and supplemented with self-reported data from the Non-Study Resources Form (NSRF). Use of non-medical resources (e.g., travel time to medical care) will also be self-reported and collected by the NSRF. The validity of self-reported data on healthcare use is well established.[96-99] Resource unit-cost information will be gathered from various relevant sources. The health-related quality-of-life preference weights required to calculate QALYs for the cost-effectiveness analysis (see Section 4.G.2) will be obtained from the EuroQol 5D (EQ-5D).[100, 101] Among generic, preference-based instruments, EQ-5D is the most widely used.[102] See Table 4 for information on the timing of data collection.

4.F.7 ANA Framework. We will use the NIH Toolbox, a brief, comprehensive, standardized assessment normed on people aged 3-85 years and administered by computer, to assess negative emotionality and executive functioning.[103, 104] The NIH Toolbox uses computerized adaptive testing, so it can be administered efficiently and accurately. Its Emotion battery, which includes measures of **negative emotionality** (e.g., anger, fear, sadness) can be completed in less than 5 minutes. The Emotion battery will be completed at baseline and then weekly during induction to accurately assess mood. Other measures will be administered at baseline only. The NIH Toolbox Cognition battery requires 30 minutes to complete and includes measures of **executive functioning,** episodic memory, language, processing speed, working memory, and attention. These comprise the Fluid Cognition Composite Score, our primary measure of executive functioning. We will supplement this battery with the Monetary Choice Questionnaire (MCQ), a delayed discounting measure from the PhenX Toolkit.[105] **Alcohol-related incentive salience** will be measured with the Obsessive-Compulsive Drinking Scale, a self-report measure of frequency and consequences of alcohol-related thoughts and behaviors,[106] and the Stimulus-Response Compatibility Task, a performance-based measure of approach-avoidance of alcohol-related cues.[107] We will administer **non-ANA measures** that might impact outcomes. These include the Stages of Change Readiness and Treatment Eagerness Scale, a measure of **motivation to change** alcohol use,[108] and **[**an item from the Treatment Experiences and Expectancies questionnaire that assesses **drinking goals.**[9]**]**

4.F.8 Adverse Events. The risk of adverse events related to study procedures is low.[109] Trained research staff will assess symptoms of alcohol withdrawal (sweating, hallucinations, disorientation, and tremors) at each visit by using the SHOT, a brief but comprehensive measure.[110] Withdrawal symptoms will be reported immediately to Dr. McDonell, and as appropriate to clinicians for treatment. All adverse events will be handled according to our Data Safety Monitoring Plan.

G. Data Analysis

4.G.1 General Data Analytic Strategy. We will analyze outcome data in a manner consistent with previous RCTs of CM.[19, 55] We will analyze our primary and secondary outcomes longitudinally across the final 12-weeks of CM (when all reinforcers are contingent on alcohol abstinence) and the 6 month follow-up period to estimate the effects of CM. Below we discuss our strategy for the primary outcome of alcohol use, which is very similar to our strategy for Aim 2 outcomes. We will compute means and standard deviations for continuous variables and percentages and frequencies for categorical variables. Measures that fail to satisfy the assumptions of parametric tests will be transformed to adjust for differences in variability or skewness. Alternatively, nonparametric tests may be used. Our alpha threshold for statistical significance will be 0.05.

4.G.2 Specific Data Analytic Techniques. Groups will be compared for differences in demographics and pre-treatment drug and alcohol use, including data collected during the baseline interview and the 4-week induction period. We will **evaluate sex as a biological variable** by investigating whether sex moderates treatment outcomes. Comparisons will use analysis of variance (ANOVA) for continuous variables and chi-square tests for categorical variables (with Bonferroni-adjusted follow-up tests if necessary). Variables that differ between groups at baseline and during the induction period will be used as covariates in subsequent analyses, as will baseline measures that co-vary with the outcome measures.

Aims 1 and 2. Recent alcohol use will be documented by using urine test results (binary outcome: EtG < 100 ng/mL or EtG > 100 ng/mL) and analyzed by using generalized estimating equations (GEE), generalized regression, and Cox proportional hazard regression techniques. Other outcome measures will be derived from recent alcohol use, such as 1) alcohol use (yes/no) over time; 2) longest duration of abstinence; 3) total number of periods of alcohol abstinence; 4) time to first alcohol abstinence; and 5) time to relapse after a specified period of alcohol abstinence. These outcomes will be examined across CM condition, time, and the CM condition-by-time interaction to determine whether one condition is superior to the others in promoting discontinuation of alcohol use. Additional analyses will be conducted by defining alcohol use in various clinically descriptive ways, such as the mean number of heavy drinking days or heavy drinking months across groups. Group comparisons of other Aim 2 outcomes will be assessed and evaluated in a similar fashion (e.g., longest duration of drug abstinence, mean HIV risk score).

Analyses will be conducted on the intent-to-treat sample (n=240). This is a conservative approach aimed at avoiding overestimation of the treatment effect. It includes all participants who were randomly assigned to treatment conditions, regardless of treatment completion. GEE models will be used to evaluate the differential change in primary and secondary outcomes by study condition over the 12-month study period (1 month before the baseline interview plus the 1-month induction, 4-month treatment, and 6-month follow-up periods), and whether condition interacts with change over time. This approach allows us to characterize between-group differences in longitudinal outcomes. We hypothesize that the modified interventions (High-Magnitude CM and Shaping CM) will lead to higher rates of alcohol abstinence than the Usual CM intervention. If the condition-by-time interaction variable is significant, we will conduct planned comparisons to examine the hypothesized pattern for each outcome. Such an approach will allow us to test this set of hypotheses with specificity. We will also conduct a direct comparison of the High-Magnitude CM group and the Shaping CM group.

Aim 3. We will conduct a systematic economic evaluation of the 3 CM conditions. We will estimate the relative economic value of each intervention by conducting cost-benefit and cost-effectiveness analyses. The economic evaluation will follow the guidelines recommended by the **Panel on Cost Effectiveness in Health and Medicine**,[5] Glick et al.[4] and Drummond.[111] Cost-benefit analyses will entail determining the costs of each condition and extrapolating the downstream savings resulting from reduced alcohol consumption and reduced alcohol-related consequences relative to Usual CM, at which point the estimated net benefit of each condition can be compared. The primary outcome will be the incremental cost-effectiveness ratio (ICER), which will be the incremental cost of one CM condition relative to another, divided by the incremental effectiveness of one CM condition relative to another. The most cost-effective condition will be identified by using the rules of dominance and extended dominance.[4, 112] The primary economic effectiveness outcome in the denominator of each ICER will be the QALY, a measure that incorporates both duration and health-related quality-of-life, and is **recommended as the primary effectiveness outcome measure** for economic evaluation studies by the Panel on Cost Effectiveness in Health and Medicine.[5] In addition to capturing a wider range of consequences than clinical effectiveness measures, QALYs permit comparisons across diseases and interventions, enabling broader economic interpretation. In addition, generally accepted cost-effectiveness thresholds for defining value have been established for QALYs, but not for clinical measures.[113-115] Nevertheless, **we will also calculate ICERs by using** the clinical effectiveness measures of **duration of alcohol abstinence**, to permit comparisons with existing economic evaluations of CM interventions that relied on measures of abstinence, and **number of heavy drinking days averted** a clinically significant alcohol outcome measure.[57]

Analyses will be conducted from the perspectives of the payer of healthcare services (i.e., the insurer), the treatment provider, and society. The payer and provider represent important perspectives, given the pivotal role played by each in sustaining the intervention, while the societal perspective is recommended as the

appropriate reference case.[5] All values will be adjusted for inflation. Since no follow-up measurements will be taken beyond 12 months following baseline, discounting for time preference will not be required.[4, 5] Given the differences in mechanisms to generate data, separate multivariable generalized-linear models (GLMs) will be estimated to predict the mean value of each resource category at each time point. GLMs permit us to choose the most appropriate mean and variance functions according to fit with the data.[4] A multivariable GLM will also be used to predict health-related quality of life preference weights for each participant at each time point. QALYs gained by the intervention will then be estimated by using the area under the curve.[4, 5, 116] The method of recycled predictions will be used to obtain the final predicted mean values for each study group and resource, which will then be summed and tested according to the relevant perspective.[4] To account for sampling uncertainty in point estimates, both the p-values and standard errors, as well as the confidence intervals for the ICER, will be estimated by using nonparametric bootstrapping techniques within the multivariable framework. Parametric methods based on parameters obtained from bootstrapping will be used to estimate acceptability curves, which illustrate the probability that the intervention is a good value for different willingness-to-pay thresholds (e.g., cost-per-QALY). Sensitivity analyses will be performed to account for uncertain precision in assumptions and parameter estimates applied in the analysis.[5]

[*Aim 4.* We will examine whether the ANA moderators (executive functioning, negative emotionality, alcohol-related incentive salience) interact with CM, and if so, with which CM condition. This will help us determine whether a given moderator significantly modified the impact of CM on alcohol abstinence. First, we will create an interaction term of CM condition by the moderator in question and examine whether the interaction is significant. Second, if the interaction is significant, we will examine whether the outcome varies within each CM condition across 3 categories of the moderator in question: 1) Mean ±1 standard deviation, 2) > 1 standard deviation, and 3) < 1 standard deviation. This will be accomplished by conducting a planned linear trend analysis or simple effects analysis for each CM condition, since we hypothesize that, for example, more executive dysfunction will produce worse treatment outcomes.[117, 118] Although we anticipate that this association will hold true for each condition, we also hypothesize that the effect will be more pronounced in the Usual CM condition. If appropriate for a clearer clinical interpretation, we will use centering (i.e., subtracting the mean from all participant values on the measure) on the moderator being examined, which is consistent with current expert recommendations.[117]**]**

Missing Data. We have developed extensive tracking procedures to reduce missing data, which will be handled in a manner consistent with current expert recommendations.[19, 69, 119-121] We will use maximum likelihood or multiple imputation and extensive sensitivity analyses, including "missing not at random" approaches, to examine the robustness of treatment effects under varying assumptions. These approaches are preferable to other data imputation approaches, such as mean imputation.[122]

4.G.3 Power Analysis. **[*Aims 1 and 2.*** Our choice of sample size for randomization (n=240) is based on statistical power calculations for the primary aim of reducing alcohol use and the secondary aim of improving other outcomes during the final 3 months of the intervention and the 6-month follow-up period. We estimated power by using **clinically significant** cutoff estimates in conjunction with previously reported effect sizes of adapted CM interventions.[6, 123] Clinically significant cutoffs derive from studies of addiction treatment clinicians, who report that a 10-12% difference in dichotomous outcomes (such as percent abstinent), or a 50% decline compared to the base rate, or a doubling of continuous measures such as number of weeks abstinent, would be enough to justify adoption of a given treatment in their practice.[6] This measure of clinical significance for continuous outcomes represents a large effect size. If used exclusively to guide our power analyses, this threshold would leave the current trial powered only on a doubling of continuous measures, which could lead to overlooking potentially meaningful effects. Thus, we powered our analysis primarily on clinically meaningful differences in binary outcomes. This approach gives us adequate power to detect clinically meaningful improvements in continuous outcomes that are less than double the continuous measures.

Background. In one sample of difficult-to-treat cocaine users, the lowest effect observed was a 10% increase in the proportion of negative urine samples submitted by **high-magnitude CM** participants compared to those in usual CM.[30] In a similar sample of cocaine users, a difference of 32% was found between usual versus high-magnitude CM groups.[32] In a sample of difficult-to-treat methadone maintained cocaine users, abstinence increased by 14% in high-magnitude CM.[31] Thus high-magnitude CM has been shown to consistently produce clinically meaningful and statistically significant results among patients who smoke or use illicit drugs.

Although **shaping CM** is less studied, a trial with hard-to-treat smokers found a 17% increase in smoking-negative samples submitted by shaping CM participants relative to usual CM participants. Shaping CM participants also submitted an average of 21 smoking-negative samples, while usual CM participants submitted an average of 3.5 smoking-negative samples. These differences meet the above criteria for clinical significance.

With regard to our **Aim 2** outcomes, we highlight findings from a study with cocaine users identified as non-responders to usual CM. When participants received high-magnitude CM, opiate abstinence increased by 14%

and benzodiazepine abstinence increased by 11% relative to when participants received usual CM.[32] In another study of CM for cocaine use in methadone patients, opiate abstinence increased by 8% in participants receiving high-magnitude CM.[31] Finally, a study utilized data from a large national schizophrenia treatment trial to establish a threshold for statistical significance in PANSS outcomes (our psychiatric severity measure). They found that a change in PANSS score of 34% was considered a clinically significant difference.[123] While this is lower than the cutoff for clinical significance noted above (i.e., 50% decline), our power analyses demonstrate that the proposed trial will be able to detect such an effect size, as well as slightly smaller differences.

Aim 1 and 2 Power Analysis. While we will randomize 240 participants, our power calculations are based on a sample size of n=180, or 60 per group, to adjust for an estimated attrition of 25% after randomization. We emphasize that our methodology for handling missing data will help to recapture some of the power lost to attrition. All power analyses use an alpha threshold for statistical significance of 0.05. Power analyses for Aims 1 and 2 are based on both binary and continuous outcome measures (e.g., negative or positive for alcohol use, longest duration of abstinence). For binary measures, our trial will have at least 85% power to detect a 7% difference between treatment groups. We will also maintain at least 83% power to detect an 8% difference in binary outcomes when conducting our planned comparisons after our initial, omnibus test. For continuous measures, our trial will have at least >99% power to detect a 50% mean reduction in measures of continuous alcohol use (e.g., heavy drinking days) between treatment groups. We will also maintain at least >99% power to detect an 50% mean decrease in continuous alcohol use measures (e.g., heavy drinking days) when conducting our planned comparisons after our initial, omnibus test. Our estimates for continuous outcomes might seem well informed but overly optimistic in the context of the literature on clinical significance. Therefore, we performed a second power analysis for clinically informative but less clinically important effects. Notably, we will still maintain at least 83% power to detect a 25% mean reduction in continuous outcome measures (e.g., heavy drinking days) between treatment groups. We will also maintain at least 88% power to detect a 30% mean reduction in continuous alcohol use measures when conducting our planned comparisons after our initial, omnibus test. This result is important, given the clinical significance threshold of a 34% decrease noted above for PANSS scores. Our power estimates for Aims 1 and 2 in this section demonstrate that we will have sufficient power to detect effects that are slightly smaller, not only than those previously reported, but also than those found to be clinically significant, even after accounting for attrition.

Aim 3 Power Analysis. Power analyses for cost-effectiveness analyses are based on assumptions regarding differences in cost and effect, standard deviations of cost and effect, the correlation of the difference in cost and effect, desired confidence and power levels, and the maximum willingness-to-pay value.[124, 125] Using a range of values for each of these factors, as well as the target sample size of 180 (60 per group), we estimated the power for the primary outcomes of these analyses: cost per QALY, cost per abstinent year, and cost per heavy drinking day averted. We estimate at least 80% power to be 95% confident that an intervention is a good value at the commonly cited lower-bound willingness-to-pay threshold of $50,000 per QALY.[114, 115] Power for the cost-benefit analysis is more speculative. However, if the intervention cost is $396 (SD=$41) and there is a cost offset of $1,000 from reductions in use of high-cost healthcare services,[2, 58] the net benefits would be $604. To achieve 80% power, the standard deviation of the cost offset would have to be less than $1,650. While difficult to forecast, it is not unreasonable to have a standard deviation less than 165% of the mean cost offset. Hence, we should have sufficient power for our cost-benefit analysis.

Aim 4 Power Analysis. For binary outcomes (e.g., submissions of EtG-negative samples), our trial will have at least 81% power to detect a 10% difference across the moderator (e.g., executive functioning) by group interactions. We will also maintain at least 80% power to detect a 13% difference in binary outcomes in our simple effects analysis or linear trend analysis to probe the interaction further. For continuous outcomes, our trial will have at least 80% power to detect a 30% mean reduction in continuous alcohol use measures (e.g., heavy drinking days) across the moderator (e.g., executive functioning) by group interactions. We will also have at least 82% power to detect a 44% mean reduction in continuous alcohol use measures (e.g., heavy drinking days) when conducting our simple effects analysis or linear trend analysis to probe the interaction further. Our planned analyses for Aim 4 provide sufficient power for nearly every scenario, even with a conservative estimate, because our missing data strategy will help to recover much of the lost power.**]**

H. Timeline

Project Phase	Year 1		Year 2		Year 3		Year 4		Year 5	
Start-up										
Recruitment										
Induction/Treatment period										
Follow-up										
Data analysis										
Dissemination of results										

PROTECTION OF HUMAN SUBJECTS & RISKS TO THE SUBJECTS

A. Overview of Human Subjects Involvement, Sources of Materials, and Risks

A.1 Human Subjects Involvement and Characteristics. Participants will be 400 heavy drinkers diagnosed with moderate to severe alcohol use disorders (AUDs) and serious mental illness (SMI): schizophrenia, schizoaffective disorder, bipolar I or II, or recurrent major depressive disorder. All participants will be receiving treatment as usual at one of two partnering treatment agencies. Participants will be involved in the study for an 11-month period. Participants who meet inclusion/exclusion criteria and provide informed consent will take part in a study baseline interview and a 4-week induction period where reinforcement is delivered for providing urine samples 3 times a week. After the 4-week induction all participants who 1) attain an average EtG score of >499 ng/mL (suggesting frequent heavy drinking) AND 2) attend at least 1 study visit during the final week of the induction period will be randomized to treatment conditions. Based on pilot data and our screening procedures which will rule out light drinkers, we anticipate at least 60% of the 400 initially eligible participants will meet criteria for randomization (n=240). These participants will then be randomized to take part in one of the three contingency management (CM) interventions. All those who do not meet criteria for randomization will be referred to other AUD treatments available at our partnering agencies, and when necessary, other local providers.

During the 16-week treatment period, participants will submit urine samples 3 times a week to assess alcohol use. The 16-week treatment period will be followed by a 6-month follow-up period where outcomes will be assessed at 1-, 3-, and 6-months after the treatment phase. During the induction and treatment periods participants will complete other study outcome measures (e.g., self-report) on a monthly basis. Participants will be reimbursed for study participation at baseline ($30), completion of the induction period ($20), and at each of their monthly interviews (8 X $20). The CM intervention duration is similar to previous studies.[55]

It is expected that the demographics of participants in this study will reflect the overall gender and ethnic characteristics of the consumers currently receiving services at Community Psychiatric Clinic and Frontier Behavioral Health.

A.2 Source of Study Materials. Research materials obtained from participants will include data on demographics, personal characteristics, psychological assessments, alcohol and drug use, urine specimens, and healthcare utilization. At the baseline interview each participant will sign a Health Insurance Portability and Accountability Act (HIPAA) form authorizing release of electronic medical records. Data and materials will be available only to authorized staff in a de-identified manner. All data with information concerning an individual participant, including urine samples and medical utilization data, will be marked with a unique code rather than the participant's name. The link between the code and the name will be stored on a password protected encrypted excel file that is only accessible by study staff. All medical utilization data will be transferred to investigators using an encrypted data transfer system. Direct identifiers will be removed from these data as soon as possible. To the extent possible, no information about a participant will be released to anyone without the participant's explicit written consent, except in the event of a medical emergency, when pertinent medical information will be given to medical personnel caring for the participant, or as required by law.

A.3 Handling of Urine Samples. Urine is not considered a biohazard. However, over the last 8 years we have handled over 20,000 urine samples and have developed detailed procedures for the handling of urine. When handling urine all study staff will wear gloves, eye protection, and lab coats. They will follow appropriate disposal procedures for all urine and laboratory consumables. If a urine sample is observed to contain blood, our research staff will immediately place the urine sample in a biohazard bag and dispose of it at appropriate Washington State University biohazard receptacles located at each site.

A.4 Potential Risks. The anticipated physical, psychological, social, and legal risks are minimal. However, as in all trials of this nature, possible outcomes include unauthorized disclosure of confidential information; discomfort or embarrassment related to urine collection or questionnaires dealing with personal habits, lifestyles, or drug or alcohol use; and possible unwanted encounters with friends or associates at the treatment center. If participants find any aspects of their involvement in the program psychologically or otherwise uncomfortable, they will be encouraged to discontinue participation and assisted in finding a more acceptable form of treatment. Alcohol withdrawal is another risk of study participation. Suicidal and homicidal thoughts and behaviors might also occur. Our team has extensive experience managing such behaviors. We describe how we will manage these risks in detail below.

B. Adequacy of Protection against Risks

B.1 Recruitment and Informed Consent. Clinic staff will give all potential participants a study brochure and a form asking if they would like to be contacted by study staff. Study staff will also describe the study at group treatment sessions and provide study contact information. Study staff will then either obtain contact

information for consumers who indicate interest in participation or will be contacted directly by interested consumers. Study recruitment materials will also be posted at locations that adults with co-occurring disorders are likely to frequent, such as transit centers and local hospitals. Study staff will contact potential participants to provide additional study information and screen for recent heavy drinking. Recruitment materials will be approved by the Washington State University Institutional Review Board. Initial interviews with potential participants and all subsequent contacts will be conducted in private offices at each study site. All information containing participant contact information will be stored in locked file cabinets in locked offices at each site.

Informed consent will be obtained in writing from all potential participants before research participation. During the initial interview, potential participants will be told about the study's purpose and all procedures involved. For those who decide to participate, study staff will administer the University of California, San Diego Brief Assessment for Capacity to Consent (UBACC), a brief screener of capacity to consent to research procedures designed for adults with psychiatric illness.[84] Those who cannot demonstrate capacity to provide informed consent on this screener will be administered the MacArthur Competency Assessment Tool for Clinical Research (MacCAT-CR), a comprehensive, gold standard tool for assessing capacity to provide informed consent for adults with SMI.[85] Only people with a demonstrated capacity to provide informed consent on these measures will be enrolled in the study. Staff will emphasize that there is no pressure to participate, that referral to alternative treatment will be provided if participants do not wish to continue in the study, and that potential participants are encouraged to speak to Dr. McDonell or study staff about any concerns related to study participation. Issues of confidentiality will be discussed in detail. Interested adults will then read and sign the informed consent form in the presence of study staff, or study staff will read the form to them and witness their signature. Two forms will be signed, one for the participant to keep and another for study records. The latter form will be kept in a locked file separate from participant data. Participants who have questions about informed consent or any other aspect of the research at any time will be directed to Dr. McDonell. Participants will complete a baseline interview within 2 weeks of signing the consent form. All research staff will receive extensive training in appropriate procedures to obtain informed consent, including administering measures of capacity to provide informed consent. They will conduct a mock consent interview for Dr. McDonell to ascertain the adequacy of the consent process and ensure consistency in the process across participants. Dr. McDonell has extensive experience in consenting adults with SMI through his previous studies.

B.2 Protection Against Other Risks. All study staff will be trained to be sensitive to issues surrounding confidentiality and other forms of participant risk. If at any time a participant expresses discomfort over any activity of the treatment program, staff will be told to discontinue the distressing activity and seek consultation to minimize risk. All data will be directly entered into a password-protected and encrypted RedCap database. The database will be hosted on a secure server at Washington State University and backed-up nightly. Study identification numbers will be the only identifier entered into the database. We will also request a Certificate of Confidentiality for this study from the NIAAA.

Potential participants will be thoroughly screened to eliminate candidates with disorders or conditions that are contraindicated, including high risk of alcohol withdrawal. Social risks from participation are expected to be minimal, since sensitive material is unlikely to be divulged. We will emphasize the confidential nature of all data collected in this study to potential participants. We will also explain thoroughly our safeguarding procedures. Participants who are judged by project investigators at any point to be a danger to self or others, or who are judged to be in grave danger due to continued alcohol or drug use or to medical or psychiatric problems, will be discontinued from the study but actively connected with appropriate treatment services. Details on this safety procedure will be included in the study consent form.

While our exclusion criteria are intended to prevent the participation of patients who are likely to experience dangerous alcohol withdrawal, we anticipate that withdrawal symptoms might occur, especially if the CM interventions are effective. Study staff will be trained to assess any and all symptoms of withdrawal and administer the SHOT, a brief and well validated measure of alcohol withdrawal symptoms, at each appointment.[110] If withdrawal symptoms are observed, they will be immediately reported to Dr. McDonell. He will consult with Dr. Ries, as well as site medical personnel to decide how to best manage these symptoms. All sites are experienced and prepared to handle alcohol withdrawal. Dr. Ries has over 30 years of experience in the medical management of alcohol withdrawal. We have used this procedure in 3 RCTs of CM for alcohol use disorders and have had no cases of dangerous alcohol withdrawal. Any event that might affect participants' safety will be immediately reported to the Dr. McDonell, who will work with local healthcare providers to ensure participant safety.

[Through their research studies and clinical experience, Drs. McDonell and Ries have extensive expertise managing suicidal and homicidal risk in adults with co-occurring disorders. Dr. McDonell, a licensed clinical psychologist, has 8 years of experience managing serious adverse events, including suicidal and violent thoughts and behavior in adults with co-occurring disorders. Through our studies, we have developed and

implemented protocols to manage suicide-risk and aggressive and violent behavior. Participants who report a suicidal attempt or self-harm behavior in the last thirty days and/or report repeated suicidal thoughts during the previous two weeks will be assessed for suicide risk by a treatment as usual clinician before they leave the clinic. The clinician will then determine the appropriate plan for managing suicide risk and refer participants to crisis or emergency services, if necessary. Study staff will also notify Dr. McDonell when suicidal behavior or ideation is reported and the outcome of the clinician's assessment. Participants will be notified of this procedure at informed consent. All research staff will be trained in this procedure by Dr. McDonell, who has over 5 years of experience providing dialectical behavior therapy to patients with chronic suicidality and addiction. We have utilized this procedure in 3 of our CM RCTs and it is effective, with no-suicidal behaviors occurring after our procedure was implemented.

Consistent with the literature, we have observed extremely low rates of homicidal ideation or violent behavior in our studies of adults with SMI. During the thousands of visits with over 300 participants, 2 participants have made verbal threats or engaged in violent behavior (shoving a staff member). We have managed each of these situations in close collaboration with treatment as usual clinicians. Because our research takes place in community clinics, research staff will be trained in and comply with all clinic safety procedures. All threats of violence will be taken seriously by study staff who will follow clinic and study procedures, staff these cases with the Drs. McDonell and Dr. Ries, and call 911 if a potentially life threatening situation occurs. Additionally, study staff will follow all laws pertaining to mandated reporting of child or elder abuse. All of these events will also be reported to the study's Data Safety Monitoring Board, the National Institutes of Health, and the Institutional Review Board, as described in the Data Safety Monitoring Plan.**]**

B.3 Adverse Events. We have developed project stopping rules to protect the safety of study participants. In the case of a Serious Adverse Event (as defined by the FDA and the Washington State University Institutional Review Board), the study will be stopped and no further enrollment will take place until an investigation of the event has been conducted by Dr. McDonell , the Co-Investigators, the Data Safety Monitoring Board, and Partnering Agency Leadership. A determination of the association of the adverse event with the study intervention will be made and appropriate modification of the protocol will be executed if an association is suspected. If protocol modifications to ensure the safety of future study participants cannot be executed, the study will be terminated. Serious Adverse Events will be reported to the Washington State University Institutional Review Board within 24 hours and to the appropriate program officer of the National Institutes of Health within 48 hours. Dr. McDonell must decide whether the event should be classified as serious or not by using standard FDA guidelines for Serious Adverse Events. Non-serious events will be reported to the Data Safety Monitoring Board quarterly and to the program official and Institutional Review Board annually.

B.4 Data Safety Monitoring Plan and Board. This Human Subjects Research meets the National Institutes of Health definition of a Phase II clinical trial. This clinical trial requires registration in ClinicalTrials.gov. Because this is a multi-site trial an appropriate Data Safety Monitoring Board and Plan has been developed and is described in the Data Safety Monitoring Plan section of the application.

C. Potential Benefits

The minimal risks anticipated for study participants are reasonable in relation to the anticipated benefits, because heavy drinkers will receive potentially valuable treatment for their heavy drinking at no cost. Identification of effective treatment of heavy drinkers who suffer from SMI may also provide useful information to current treatment systems and thereby benefit others. The study will also provide useful data about the cost effectiveness of CM interventions to the partnering agencies, as well as information to clinicians about CM adaptations that might be particularly effective for specific individuals. The data collected in the proposed study will have a high level of external validity and may serve to render CM interventions more effective for this population.

D. Importance of Knowledge to be Gained

The results of the proposed research may lead to more effective treatment for heavy drinkers who suffer from SMI and heavy drinkers more generally. Potential improvements in secondary outcomes, estimates of cost-effectiveness, and a better understanding of the moderators of alcohol abstinence are important contributions to the scientific literature and will provide important information to consumers, treatment providers and policy makers. Once tested, this intervention can be disseminated in community mental health centers and other places where adults with alcohol use disorders and SMI are treated.

E. ClinicalTrials.gov Requirements

This proposal includes an applicable clinical trial that requires registration in ClinicalTrials.gov.

LITERATURE CITED

1. Murphy SM, Campbell AN, Ghitza UE, Kyle TL, Bailey GL, Nunes EV, Polsky D. Cost-effectiveness of an internet-delivered treatment for substance abuse: Data from a multisite randomized controlled trial. *Drug Alcohol Depend.* 2016;161:119-126.
2. Polsky D, Glick HA, Yang J, Subramaniam GA, Poole SA, Woody GE. Cost-effectiveness of extended buprenorphine–naloxone treatment for opioid-dependent youth: data from a randomized trial. *Addiction.* 2010;105(9):1616-1624.
3. Drummond MF, Davies L. Economic analysis alongside clinical trials. *Int J Technol Assess Health Care.* 1991;7(4):561-573.
4. Glick HA, Doshi JA, Sonnad SS, Polsky D. *Economic evaluation in clinical trials*: Oxford University Press; 2014.
5. Gold MR, Siegel JE, Russell LB, Weinstein MC. *Cost-effectiveness in health and medicine*. New York: Oxford University Press; 1996.
6. Miller WR, Manuel JK. How large must a treatment effect be before it matters to practioners? An estimation method and demonstration. *Drug Alcohol Rev.* 2008;27:524-528.
7. Srebnik D, Sugar A, Coblentz P, McDonell MG, Angelo F, Lowe JM, Ries RK, Roll J. Acceptability of contingency management among clinicians and clients within a co-occurring mental health and substance use treatment program. *Am J Addict.* 2013;22(5):432-436.
8. Kirby KC, Benishek LA, Dugosh KL, Kerwin ME. Substance abuse treatment providers' beliefs and objections regarding contingency management: implications for dissemination. *Drug Alcohol Depend.* 2006;85(1):19-27.
9. Dunn KE, Strain EC. Pretreatment alcohol drinking goals are associated with treatment outcomes. *Alcohol Clin Exp Res.* 2013;37(10):1745-1752.
10. Regier DA, Farmer ME, Rae DS, Locke BZ, Keith SJ, Judd LL, Goodwin FK. Comorbidity of mental disorders with alcohol and other drug abuse. Results from the Epidemiologic Catchment Area (ECA) Study. *JAMA.* 1990;264(19):2511-2518.
11. Gonzalez VM, Bradizza CM, Vincent PC, Stasiewicz PR, Paas ND. Do individuals with a severe mental illness experience greater alcohol and drug-related problems? A test of the supersensitivity hypothesis. *Addict Behav.* 2007;32(3):477-490.
12. Dixon L. Dual diagnosis of substance abuse in schizophrenia: prevalence and impact on outcomes. *Schizophr Res.* 1999;35 Suppl:S93-100.
13. DHHS. *Substance Abuse Treatment for Persons with Co-occurring Disorders. A Treatment Improvement Protocol TIP 42*. Rockville MD: US DHHS; 2005.
14. Buckley PF, Brown ES. Prevalence and consequences of dual diagnosis. *J Clin Psychiatry.* 2006;67(7):e01.
15. Smelson DA, Tunis SL, Nyhuis AW, Faries DE, Kinon BJ, Ascher-Svanum H. Antipsychotic treatment discontinuation among individuals with schizophrenia and co-occurring substance use. *J Clin Psychopharmacol.* 2006;26(6):666-667.
16. Kreyenbuhl J, Nossel IR, Dixon LB. Disengagement from mental health treatment among individuals with schizophrenia and strategies for facilitating connections to care: a review of the literature. *Schizophr Bull.* 2009;35(4):696-703.
17. Bradizza CM, Stasiewicz PR, Dermen KH. Behavioral interventions for individuals dually-diagnosed with a severe mental illness and a substance use disorder. *Curr Addict Rep.* 2014;1(4):243-250.
18. Hunt GE, Siegfried N, Morley K, Sitharthan T, Cleary M. Psychosocial interventions for people with both severe mental illness and substance misuse. *Cochrane Database Syst Rev.* 2013;10:CD001088.
19. McDonell MG, Srebnik D, Angelo F, McPherson S, Lowe JM, Sugar A, Short RA, Roll JM, Ries RK. Randomized controlled trial of contingency management for stimulant use in community mental health patients with serious mental illness. *Am J Psychiatry.* 2013;170(1):94-101.
20. Leickly E, McDonell MG, Vilardaga R, Angelo FA, Lowe JM, McPherson S, Srebnik D, Roll JM, Ries RK. High levels of agreement between clinic-based ethyl glucuronide (EtG) immunoassays and laboratory-based mass spectrometry. *Am J Drug Alcohol Abuse.* 2015;41(3):246-250.
21. Lowe J, McDonell M, Leickly E, Angelo F, Vilardaga R, McPherson S, Srebnik D, Roll J, Ries R. Determining ethyl glucuronide cutoffs when detecting self-reported alcohol use in addiction treatment patients. *Alcohol Clin Exp Res.* 2015;39(5):905-910.

22. McDonell MG, Skalisky J, Leickly E, McPherson S, Battalio S, Nepom JR, Srebnik D, Roll J, Ries RK. Using ethyl glucuronide in urine to detect light and heavy drinking in alcohol dependent outpatients. *Drug Alcohol Depend.* 2015;157:184-187.

23. McDonell MG, Srebnik D, Angelo F, Sugar AM, Howell D, Rainey C, Roll J, Short R, Ries R. Evaluation of ethyl glucuronide immunoassay urinalysis in five alcohol-dependent outpatients. *Am J Addict.* 2011;20(5):482-484.

24. McDonell MG, Howell DN, McPherson S, Cameron JM, Srebnik D, Roll JM, Ries RK. Voucher-based reinforcement for alcohol abstinence using the ethyl-glucuronide alcohol biomarker. *J Appl Behav Anal.* 2012;45(1):161-165.

25. McDonell M, Leickly E, McPherson S, Skalisky J, Srebnik D, Angelo F, Vilardaga R, Nepom J, Roll J, Ries R. A randomized controlled trial of ethyl glucuronide-based contingency management for outpatients with co-occurring alcohol use disorders and serious mental illness. *Am J Psychiatry.* In Press.

26. Lamb RJ, Morral AR, Galbicka G, Kirby KC, Iguchi MY. Shaping reduced smoking in smokers without cessation plans. *Exp Clin Psychopharmacol.* 2005;13(2):83-92.

27. Preston KL, Umbricht A, Wong CJ, Epstein DH. Shaping cocaine abstinence by successive approximation. *J Consult Clin Psychol.* 2001;69(4):643-654.

28. Lamb RJ, Morral AR, Kirby KC, Iguchi MY, Galbicka G. Shaping smoking cessation using percentile schedules. *Drug Alcohol Depend.* 2004;76(3):247-259.

29. Lamb RJ, Kirby KC, Morral AR, Galbicka G, Iguchi MY. Shaping smoking cessation in hard-to-treat smokers. *J Consult Clin Psychol.* 2010;78(1):62-71.

30. Petry NM, Barry D, Alessi SM, Rounsaville BJ, Carroll KM. A randomized trial adapting contingency management targets based on initial abstinence status of cocaine-dependent patients. *J Consult Clin Psychol.* 2012;80(2):276-285.

31. Dallery J, Silverman K, Chutuape MA, Bigelow GE, Stitzer ML. Voucher-based reinforcement of opiate plus cocaine abstinence in treatment-resistant methadone patients: effects of reinforcer magnitude. *Exp Clin Psychopharmacol.* 2001;9(3):317-325.

32. Silverman K, Chutuape MA, Bigelow GE, Stitzer ML. Voucher-based reinforcement of cocaine abstinence in treatment-resistant methadone patients: effects of reinforcement magnitude. *Psychopharmacology (Berl).* 1999;146(2):128-138.

33. Kwako LE, Momenan R, Litten RZ, Koob GF, Goldman D. Addictions Neuroclinical Assessment: A neuroscience-based framework for addictive disorders. *Biol Psychiatry.* 2016;80(3):179-189.

34. Koskinen J, Lohonen J, Koponen H, Isohanni M, Miettunen J. Prevalence of alcohol use disorders in schizophrenia--a systematic review and meta-analysis. *Acta Psychiatr Scand.* 2009;120(2):85-96.

35. Bauer MS, Altshuler L, Evans DR, Beresford T, Williford WO, Hauger R, Team VACS. Prevalence and distinct correlates of anxiety, substance, and combined comorbidity in a multi-site public sector sample with bipolar disorder. *J Affect Disord.* 2005;85(3):301-315.

36. Davis L, Uezato A, Newell JM, Frazier E. Major depression and comorbid substance use disorders. *Curr Opin Psychiatry.* 2008;21(1):14-18.

37. Jackson CT, Fein D, Essock SM, Mueser KT. The effects of cognitive impairment and substance abuse on psychiatric hospitalizations. *Community Ment Health J.* 2001;37(4):303-312.

38. Buckley PF. Prevalence and consequences of the dual diagnosis of substance abuse and severe mental illness. *J Clin Psychiatry.* 2006;67 Suppl 7:5-9.

39. Galanter M, Dermatis H, Egelko S, De Leon G. Homelessness and mental illness in a professional- and peer-led cocaine treatment clinic. *Psychiatr Serv.* 1998;49(4):533-535.

40. Clark RE, Samnaliev M, McGovern MP. Impact of substance disorders on medical expenditures for medicaid beneficiaries with behavioral health disorders. *Psychiatr Serv.* 2009;60(1):35-42.

41. Bennett ME, Bellack AS, Gearon JS. Treating substance abuse in schizophrenia. An initial report. *J Subst Abuse Treat.* 2001;20(2):163-175.

42. Watkins KE, Burnam A, Kung FY, Paddock S. A national survey of care for persons with co-occurring mental and substance use disorders. *Psychiatr Serv.* 2001;52(8):1062-1068.

43. Epstein JF, Hourani LL, Heller DC. Predictors of treatment receipt among adults with a drug use disorder. *Am J Drug Alcohol Abuse.* 2004;30(4):841-869.

44. Barrowclough C, Haddock G, Wykes T, Beardmore R, Conrod P, Craig T, Davies L, Dunn G, Eisner E, Lewis S, Moring J, Steel C, Tarrier N. Integrated motivational interviewing and cognitive behavioural

therapy for people with psychosis and comorbid substance misuse: randomised controlled trial. *BMJ.* 2010;341:c6325.
45. Baker A, Bucci S, Lewin TJ, Kay-Lambkin F, Constable PM, Carr VJ. Cognitive-behavioural therapy for substance use disorders in people with psychotic disorders: Randomised controlled trial. *Br J Psychiatry.* 2006;188:439-448.
46. Weiss RD, Griffin ML, Jaffee WB, Bender RE, Graff FS, Gallop RJ, Fitzmaurice GM. A "community-friendly" version of integrated group therapy for patients with bipolar disorder and substance dependence: a randomized controlled trial. *Drug Alcohol Depend.* 2009;104(3):212-219.
47. Drake RE, McHugo GJ, Clark RE, Teague GB, Xie H, Miles K, Ackerson TH. Assertive community treatment for patients with co-occurring severe mental illness and substance use disorder: a clinical trial. *Am J Orthopsychiatry.* 1998;68(2):201-215.
48. Bellack AS, Bennett ME, Gearon JS, Brown CH, Yang Y. A randomized clinical trial of a new behavioral treatment for drug abuse in people with severe and persistent mental illness. *Arch Gen Psychiatry.* 2006;63(4):426-432.
49. McFarland BH, Gabriel RM. Datapoints: Service availability for persons with co-occurring conditions. *Psychiatr Serv.* 2004;55(9):978.
50. Wakefield PA, Randall GE, Richards DA. Identifying barriers to mental health system improvements: An examination of community participation in Assertive Community Treatment programs. *Int J Ment Health Syst.* 2011;5(1):27.
51. Brunette MF, Asher D, Whitley R, Lutz WJ, Wieder BL, Jones AM, McHugo GJ. Implementation of integrated dual disorders treatment: a qualitative analysis of facilitators and barriers. *Psychiatr Serv.* 2008;59(9):989-995.
52. Lussier JP, Heil SH, Mongeon JA, Badger GJ, Higgins ST. A meta-analysis of voucher-based reinforcement therapy for substance use disorders. *Addiction.* 2006;101(2):192-203.
53. Prendergast M, Podus D, Finney J, Greenwell L, Roll J. Contingency management for treatment of substance use disorders: a meta-analysis. *Addiction.* 2006;101(11):1546-1560.
54. Dutra L, Stathopoulou G, Basden SL, Leyro TM, Powers MB, Otto MWCIN. A meta-analytic review of psychosocial interventions for substance use disorders. *Am J Psychiatry.* 2008;165(2):179-187.
55. Benishek LA, Dugosh KL, Kirby KC, Matejkowski J, Clements NT, Seymour BL, Festinger DS. Prize-based contingency management for the treatment of substance abusers: a meta-analysis. *Addiction.* 2014;109(9):1426-1436.
56. Griffith JD, Rowan-Szal GA, Roark RR, Simpson DD. Contingency management in outpatient methadone treatment: a meta-analysis. *Drug Alcohol Depend.* 2000;58(1-2):55-66.
57. Shearer J, Tie H, Byford S. Economic evaluations of contingency management in illicit drug misuse programmes: A systematic review. *Drug Alcohol Rev.* 2015;34(3):289-298.
58. Murphy SM, McDonell MG, McPherson S, Srebnik D, Angelo F, Roll JM, Ries RK. An economic evaluation of a contingency-management intervention for stimulant use among community mental health patients with serious mental illness. *Drug Alcohol Depend.* 2015;153:293-299.
59. Lopez-Nunez C, Alonso-Perez F, Pedrosa I, Secades-Villa R. Cost-effectiveness of a voucher-based intervention for smoking cessation. *Am J Drug Alcohol Abuse.* 2016;42(3):296-305.
60. Sindelar J, Elbel B, Petry NM. What do we get for our money? Cost-effectiveness of adding contingency management. *Addiction.* 2007;102(2):309-316.
61. Rash CJ, Petry NM, Kirby KC, Martino S, Roll J, Stitzer ML. Identifying provider beliefs related to contingency management adoption using the contingency management beliefs questionnaire. *Drug Alcohol Depend.* 2012;121(3):205-212.
62. Administration VH. VHA Programs for Veterans with Substance Use Disorders (SUD). In: Administration VH, ed: Veterans Health Administration, Washington, DC; 2011.
63. Ries RK, Dyck DG, Short R, Srebnik D, Fisher A, Comtois KA. Outcomes of managing disability benefits among patients with substance dependence and severe mental illness. *Psychiatr Serv.* 2004;55(4):445-447.
64. McDonell M, McPherson S, Vilardaga R, Srebnik D, Angelo FN, Leickly E, Saxon AJ, Roll J, Ries R. Preliminary findings: Contingency management targeting psycho-stimulant use results in secondary decreases in smoking for severely mentally ill adults. *Am J Addict.* 2014;23(4):407-410.
65. Angelo FN, McDonell MG, Lewin MR, Srebnik D, Lowe J, Roll J, Ries R. Predictors of stimulant abuse treatment outcomes in severely mentally ill outpatients. *Drug Alcohol Depend.* 2013;131(1-2):162-165.

66. Petry NM, Martin B, Cooney JL, Kranzler HR. Give them prizes, and they will come: contingency management for treatment of alcohol dependence. *J Consult Clin Psychol.* 2000;68(2):250-257.

67. Cunningham C, Stitzer M, Campbell AN, Pavlicova M, Hu MC, Nunes EV. Contingency management abstinence incentives: Cost and implications for treatment tailoring. *J Subst Abuse Treat.* 2015.

68. Stitzer ML, Peirce J, Petry NM, Kirby K, Roll J, Krasnansky J, Cohen A, Blaine J, Vandrey R, Kolodner K, Li R. Abstinence-based incentives in methadone maintenance: Interaction with intake stimulant test results. *Exp Clin Psychopharmacol.* 2007;15(4):344-350.

69. McPherson S, Packer RR, Cameron JM, Howell DN, Roll JM. Biochemical marker of use is a better predictor of outcomes than self-report metrics in a contingency management smoking cessation analog study. *Am J Addict.* 2014;23(1):15-20.

70. Romanowich P, Lamb RJ. The effects of percentile versus fixed criterion schedules on smoking with equal incentive magnitude for initial abstinence. *Exp Clin Psychopharmacol.* 2014;22(4):348-355.

71. Packer RR, Howell DN, McPherson S, Roll JM. Investigating reinforcer magnitude and reinforcer delay: a contingency management analog study. *Exp Clin Psychopharmacol.* 2012;20(4):287-292.

72. Carroll KM, Ball SA, Nich C, O'Connor PG, Eagan DA, Frankforter TL, Triffleman EG, Shi J, Rounsaville BJ. Targeting behavioral therapies to enhance naltrexone treatment of opioid dependence: efficacy of contingency management and significant other involvement. *Arch Gen Psychiatry.* 2001;58(8):755-761.

73. Petry NM, Alessi SM, Barry D, Carroll KM. Standard magnitude prize reinforcers can be as efficacious as larger magnitude reinforcers in cocaine-dependent methadone patients. *J Consult Clin Psychol.* 2015;83(3):464-472.

74. Rohsenow DJ, Tidey JW, Kahler CW, Martin RA, Colby SM, Sirota AD. Intolerance for withdrawal discomfort and motivation predict voucher-based smoking treatment outcomes for smokers with substance use disorders. *Addict Behav.* 2015;43:18-24.

75. Petry NM, Rash CJ, Easton CJ. Contingency management treatment in substance abusers with and without legal problems. *J Am Acad Psychiatry Law.* 2011;39(3):370-378.

76. Schepis TS, Cavallo DA, Kong G, Liss T, Liss A, Krishnan-Sarin S. Predicting initiation of smoking cessation treatment and outcome among adolescents using stressful life events and coping style. *Subst Abus.* 2015;36(4):478-485.

77. Ford JD, Hawke J, Alessi S, Ledgerwood D, Petry N. Psychological trauma and PTSD symptoms as predictors of substance dependence treatment outcomes. *Behav Res Ther.* 2007;45(10):2417-2431.

78. Krishnan-Sarin S, Reynolds B, Duhig AM, Smith A, Liss T, McFetridge A, Cavallo DA, Carroll KM, Potenza MN. Behavioral impulsivity predicts treatment outcome in a smoking cessation program for adolescent smokers. *Drug Alcohol Depend.* 2007;88(1):79-82.

79. Manning V, Betteridge S, Wanigaratne S, Best D, Strang J, Gossop M. Cognitive impairment in dual diagnosis inpatients with schizophrenia and alcohol use disorder. *Schizophr Res.* 2009;114(1-3):98-104.

80. Ames SL, Wong SW, Bechara A, Cappelli C, Dust M, Grenard JL, Stacy AW. Neural correlates of a Go/NoGo task with alcohol stimuli in light and heavy young drinkers. *Behav Brain Res.* 2014;274:382-389.

81. Bigelow GE, Silverman K. Theoretical foundations of contingency management treatments for drug abuse. In: Higgins ST, Silverman K, eds. *Motivating Behavior Change among Illicit-drug Users: Research on Contingency Management Interventions*. Washington DC: American Psychological Association; 1999:15-31.

82. van Amsterdam J, van den Brink W. Reduced-risk drinking as a viable treatment goal in problematic alcohol use and alcohol dependence. *J Psychopharmacol.* 2013;27(11):987-997.

83. Ambrogne JA. Reduced-risk drinking as a treatment goal: what clinicians need to know. *J Subst Abuse Treat.* 2002;22(1):45-53.

84. Jeste DV, Palmer BW, Appelbaum PS, Golshan S, Glorioso D, Dunn LB, Kim K, Meeks T, Kraemer HC. A new brief instrument for assessing decisional capacity for clinical research. *Arch Gen Psychiatry.* 2007;64(8):966-974.

85. Appelbaum P, Grisso T. *MacCAT-CR: MacArthur Competence Assessment Tool for Clinical Research.* Sarasota, FL: Professional Resource Press; 2001.

86. APA. *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*. Arlington, VA: American Psychiatric Association; 2013.

87. Silverman K, Wong CJ, Needham M, Diemer KN, Knealing T, Crone-Todd D, Fingerhood M, Nuzzo P, Kolodner K. A randomized trial of employment-based reinforcement of cocaine abstinence in injection drug users. *J Appl Behav Anal.* 2007;40(3):387-410.
88. Roll JM, Higgins ST. A within-subject comparison of three different schedules of reinforcement of drug abstinence using cigarette smoking as an exemplar. *Drug Alcohol Depend.* 2000;58(1-2):103-109.
89. First M, Williams J, Karg R, Spitzer R. *Structured Clinical Interview for DSM-5 Disorders, Clinical Trials Version (SCID-5-CT)*. Arlington, VA: American Psychiatric Association; 2015.
90. Sobell LC, Sobell MB. Alcohol timeline followback (TLFB). *Handbook of Psychiatric Measures*. Washington DC: American Psychiatric Association; 2000:477-479.
91. McLellan AT, Kushner H, Metzger D, Peters R, Smith I, Grissom G, Pettinati H, Argeriou M. The Fifth Edition of the Addiction Severity Index. *J Subst Abuse Treat.* 1992;9(3):199-213.
92. Heatherton TF, Kozlowski LT, Frecker RC, Fagerstrom KO. The Fagerstrom Test for Nicotine Dependence: A revision of the Fagerstrom Tolerance Questionnaire. *Br J Addict.* 1991;86(9):1119-1127.
93. Darke S, Hall W, Heather N, Ward J, Wodak A. The reliability and validity of a scale to measure HIV risk-taking behaviour among intravenous drug users. *AIDS.* 1991;5(2):181-185.
94. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull.* 1987;13(2):261-276.
95. French MT, Dunlap LJ, Zarkin GA, McGeary KA, McLellan AT. A structured instrument for estimating the economic cost of drug abuse treatment. The Drug Abuse Treatment Cost Analysis Program (DATCAP). *J Subst Abuse Treat.* 1997;14(5):445-455.
96. Brown JB, Adams ME. Patients as reliable reporters of medical care process. Recall of ambulatory encounter events. *Med Care.* 1992;30(5):400-411.
97. Harlow SD, Linet MS. Agreement between questionnaire data and medical records. The evidence for accuracy of recall. *Am J Epidemiol.* 1989;129(2):233-248.
98. Roberts RO, Bergstralh EJ, Schmidt L, Jacobsen SJ. Comparison of self-reported and medical record health care utilization measures. *J Clin Epidemiol.* 1996;49(9):989-995.
99. Wallihan DB, Stump TE, Callahan CM. Accuracy of self-reported health services use and patterns of care among urban older adults. *Med Care.* 1999;37(7):662-670.
100. Dolan P. Modeling valuations for EuroQol health states. *Med Care.* 1997;35(11):1095-1108.
101. EuroQol G. EuroQol--a new facility for the measurement of health-related quality of life. *Health Policy.* 1990;16(3):199-208.
102. Richardson J, Mckie J, Bariola E. Multiattribute utility instruments and their use. In: Culyer AJ, ed. *Encylopedia of Health Economics*: Elsevier; 2014:341-357.
103. Weintraub S, Dikmen SS, Heaton RK, Tulsky DS, Zelazo PD, Bauer PJ, Carlozzi NE, Slotkin J, Blitz D, Wallner-Allen K, Fox NA, Beaumont JL, Mungas D, Nowinski CJ, Richler J, Deocampo JA, Anderson JE, Manly JJ, Borosh B, Havlik R, Conway K, Edwards E, Freund L, King JW, Moy C, Witt E, Gershon RC. Cognition assessment using the NIH Toolbox. *Neurology.* 2013;80(11 Suppl 3):S54-64.
104. Salsman JM, Butt Z, Pilkonis PA, Cyranowski JM, Zill N, Hendrie HC, Kupst MJ, Kelly MA, Bode RK, Choi SW, Lai JS, Griffith JW, Stoney CM, Brouwers P, Knox SS, Cella D. Emotion assessment using the NIH Toolbox. *Neurology.* 2013;80(11 Suppl 3):S76-86.
105. Kirby KN, Petry NM, Bickel WK. Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls. *J Exp Psychol Gen.* 1999;128(1):78-87.
106. Nakovics H, Diehl A, Croissant B, Mann K. Modifications of the Obsessive Compulsive Drinking Scale (OCDS-G) for use in longitudinal studies. *Addict Behav.* 2008;33(10):1276-1281.
107. Barkby H, Dickson JM, Roper L, Field M. To approach or avoid alcohol? Automatic and self-reported motivational tendencies in alcohol dependence. *Alcohol Clin Exp Res.* 2012;36(2):361-368.
108. Miller WR, Tonigan JS. Assessing drinkers' motivation for change: The stages of change readiness and treatment eagerness scale (SOCRATES). *Psychol Addict Behav.* 1996;20(2):81-89.
109. Petry NM, Roll JM, Rounsaville BJ, Ball SA, Stitzer M, Peirce JM, Blaine J, Kirby KC, McCarty D, Carroll KM. Serious adverse events in randomized psychosocial treatment studies: Safety or arbitrary edicts? *J Consult Clin Psychol.* 2008;76(6):1076-1082.
110. Gray S, Borgundvaag B, Sirvastava A, Randall I, Kahan M. Feasibility and reliability of the SHOT: A short scale for measuring pretreatment severity of alcohol withdrawal in the emergency department. *Acad Emerg Med.* 2010;17(10):1048-1054.

111. Drummond MF. *Methods for the economic evaluation of health care programmes*: Oxford university press; 2005.

112. Johannesson M, Weinstein MC. On the decision rules of cost-effectiveness analysis. *J Health Econ.* 1993;12(4):459-467.

113. Hirth RA, Chernew ME, Miller E, Fendrick AM, Weissert WG. Willingness to pay for a quality-adjusted life year: in search of a standard. *Med Decis Making.* 2000;20(3):332-342.

114. Braithwaite RS, Meltzer DO, King JT, Jr., Leslie D, Roberts MS. What does the value of modern medicine say about the $50,000 per quality-adjusted life-year decision rule? *Med Care.* 2008;46(4):349-356.

115. Neumann PJ, Cohen JT, Weinstein MC. Updating cost-effectiveness--the curious resilience of the $50,000-per-QALY threshold. *N Engl J Med.* 2014;371(9):796-797.

116. Matthews JN, Altman DG, Campbell MJ, Royston P. Analysis of serial measurements in medical research. *BMJ.* 1990;300(6719):230-235.

117. Aiken L, West S. *Multiple regression: Testing and Interpreting Interactions*. Newbury Park, CA: Sage; 1991.

118. MacKinnon DP, Luecken LJ. How and for whom? Mediation and moderation in health psychology. *Health Psychol.* 2008;27(2S):S99-S100.

119. Graham JW. Missing data analysis: Making it work in the real world. *Annu Rev Psychol.* 2009;60:549-576.

120. Enders CK. Missing not at random models for latent growth curve analyses. *Psychol Methods.* 2011;16(1):1-16.

121. McPherson S, Barbosa-Leiker C, McDonell M, Howell D, Roll J. Longitudinal missing data strategies for substance use clinical trials using generalized estimating equations: An example with a buprenorphine trial. *Hum Psychopharmacol.* 2013;28(5):506-515.

122. Enders CK. *Applied Missing Data Analysis*. New York: Guilford Press; 2010.

123. Hermes ED, Sokoloff D, Stroup TS, Rosenheck RA. Minimum clinically important difference in the Positive and Negative Syndrome Scale with data from the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE). *J Clin Psychiatry.* 2012;73(4):526-532.

124. Glick HA. Sample size and power for cost-effectiveness analysis (part 1). *Pharmacoeconomics.* 2011;29(3):189-198.

125. Glick HA. Sample Size and Power for Cost-Effectiveness Analysis (Part 2). *Pharmacoeconomics.* 2011;29(4):287-296.