Evaluation of Contingency Management (CM) for Substance Use Disorder (SUD)

Reviewers: Rory A. Pfund, Meredith K. Ginley, Cassandra L. Boness, Kristyn Zajac, Carla J. Rash, Katie Witkiewitz
Dates of Review: July 25, 2021 to December 13, 2022
Online Recommendation: https://div12.org/treatment/prize-based-contingency-management-for-mixed-substance-abuse-dependence/
Recommended Citation: Pfund, R. A., Ginley, M. K., Boness, C. L., Zajac, K., Rash, C. J., & Witkiewitz, K. (2021, December 22). Evaluation of Contingency Management (CM) for Substance Use Disorder (SUD). Retrieved from osf.io/zme9n

1. Treatment Nomination and Committee Formation

The following document outlines the evaluation of CM as an empirically supported treatment for SUD according to the Tolin et al. (2015) criteria. A formal letter of intent to evaluate CM for SUD was submitted on July 16, 2021, and the Committee on Science and Practice officially approved the evaluation on July 25, 2021. This evaluation is part of a larger effort to revise and update ratings on the strength of evidence for psychological treatments for given diagnoses.

The committee was formed of six members: Cassandra L. Boness, PhD (Center on Alcohol, Substance use, And Addictions [CASAA]), Meredith K. Ginley, PhD (East Tennessee State University), Rory A. Pfund, PhD (CASAA), Carla J. Rash, PhD (University of Connecticut School of Medicine), Katie Witkiewitz, PhD (Center on Alcohol, Substance use, And Addictions, and Kristyn Zajac, PhD (University of Connecticut School of Medicine). Four members members were content experts in CM for SUD (Drs. Pfund, Ginley, Rash, and Zajac), three were experts in systematic review and meta-analysis (Drs. Pfund, Ginley, and Boness), and two were experts in treatment for SUD without allegiance to CM (Drs. Boness and Witkiewitz).

Conflicts of interest were as followed:
1. Drs. Pfund, Ginley, Rash, and Zajac were authors on two of the five meta-analyses included in the present evaluation of CM.
2. Dr. Boness was a member of the Division 12 Committee on Science and Practice and aided in the development of the manual used to guide this evaluation.

No other conflicts of interest were declared.

2. Locating and Screening Reviews for Inclusion

Consistent with Tolin et al. (2015), the current synthesis was limited to meta-analyses published in the past two years. Eight databases were searched, including Academic Search Complete, Cochrane Database of Systematic Reviews, Embase, Google Scholar, PsycINFO, PubMed, Science Direct, and Web of Science Core Collection. Search terms included the following: TI (“contingency management” OR “voucher” OR “prize” OR “behavioral contracting” OR “token economy” OR “motivational incentives” OR “incentives”) AND (TI ((“review” OR “systematic review” OR “quantitative review” OR “meta analysis” OR “meta-analysis”) OR SU (“review” OR “systematic review” OR “quantitative review” OR “meta analysis” OR “meta-analysis”)) NOT TI (“qualitative review” OR “narrative review”) where TI = title and SU = subject. There were no restrictions on language or type of publication. Initial search results returned 206 records. The search was conducted on September 9, 2021.

Meta-analyses met inclusion criteria if they focused on CM for SUD among individuals aged 18 years or older. Exclusion criteria included lack of reporting on cannabis, stimulant, and/or opioids outcomes.

Reviewers double coded each of the 206 returned records as eligible, not eligible, or possibly eligible based on their title and abstract. For records coded as eligible or possibly eligible, full texts were obtained and read to further determine eligibility. All discrepancies were resolved via consensus among the two reviews. Of the 33 eligible records, 24 were excluded because they were not a quantitative review, and 4 were excluded because they did not report cannabis, stimulant, and/or opioid outcomes. A total of 5 meta-analyses were deemed eligible for inclusion in the current review (see Figure 1).

For each of the 5 eligible meta-analyses, committee members double coded the PICOTS (population, intervention, comparison, outcomes, timeline, setting) criteria (e.g., Schardt et al., 2007), and discrepancies were resolved through consensus. PICOTS allows for a full consideration of review characteristics (see Table 1) and assists the reader in evaluating the appropriateness of the eligible reviews included for answering the clinical question of interest.

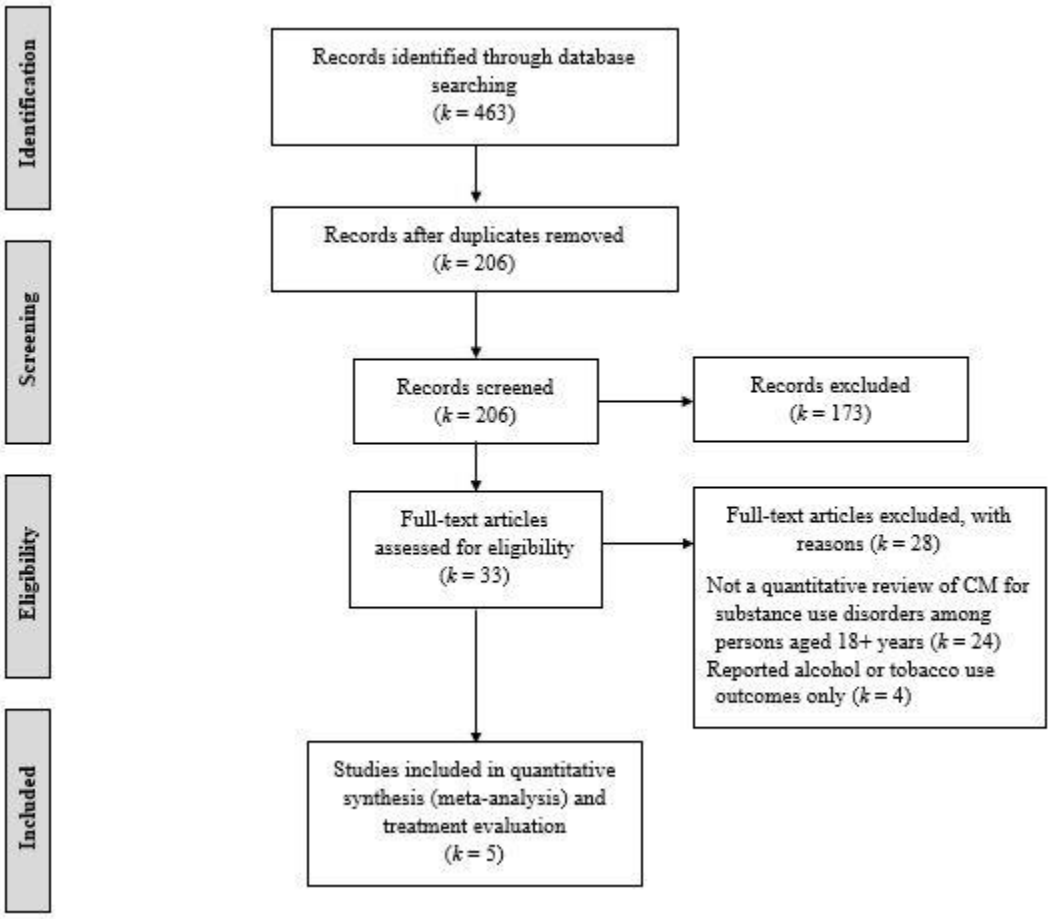

Figure 1. Flowchart for CM Search Process. This figure illustrates the search process for locating reviews eligible for inclusion in the treatment evaluation of CM for SUD. CM = contingency management; SUD = substance use disorder; *k* = number of unique records.

Table 1. Meta-Analyses Used in the Present Evaluation of CM as an Empirically Supported Treatment for SUD

Meta-Analysis	# of studies	Population	Intervention	Comparison condition	Outcome(s)	Timepoints	Setting
Bolívar et al. (2021)	59	Individuals with SUD	CM	Placebo; TAU Inactive control	Abstinence; attendance; cigarette smoking; medication use	Posttreatment; Follow up (1-12 weeks)	Medical settings
Destoop et al. (2021)	4	Individuals with concurrent SUD and psychotic disorder	CM	Placebo; TAU; Inactive control	Abstinence	Posttreatment; Follow up (12-84 weeks)	NS
Ginley et al. (2021)	23	Individuals with SUD	CM	Active control; Placebo; TAU; Inactive control	Abstinence	Follow up (6-52 weeks)	Community-based outpatient clinics; Medical settings
Pfund et al. (2022)	10	Individuals with SUD	CM	Inactive control	Abstinence; attendance	Posttreatment	NS
Sheridan Rains et al. (2020)	8	Individuals with SUD	CM	Active control; TAU	Abstinence	Posttreatment	NS

Note. Posttreatment outcomes were assessed at treatment end or throughout treatment. CM = contingency management; NS = not specified; SUD = substance use disorder; TAU = treatment as usual.

3. Assessment of Review Quality

3.1 AMSTAR2 Quality Ratings Across Eligible Studies

For each of the five included meta-analyses, committee members triple coded the AMSTAR2 (Shea et al., 2017) items, and discrepancies among coders were resolved by consensus. Consistent with the evaluation of the Tolin et al. (2015) criteria for cognitive-behavioral therapy for insomnia (Boness et al., 2020), we deemed the same AMSTAR2 items as critical: (a) including components of PICOTS in the research questions and inclusion criteria, (b) using a comprehensive search strategy, (c) describing the included studies in adequate detail, (d) using appropriate methods for statistical combination of results, (e) accounting for risk of bias in individual studies, and (f) providing explanation for a discussion for any heterogeneity observed in the results.

Table 2

AMSTAR2 Results for Eligible Studies

Item	Bolívar et al. (2021)	Destoop et al. (2021)	Ginley et al. (2021)	Pfund et al. (2022)	Sheridan Rains et al. (2020)
1. Did the research questions and inclusion criteria for the review include components of PICO?	Y	Y	Y	Y	Y
2. Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?	Y	N	N	N	Y
3. Did the review authors explain their selection of the study designs for inclusion in the review?	N	N	Y	Y	Y
4. Did the review authors use a comprehensive literature search strategy?	N	N	N	N	PY
5. Did the review authors perform study selection in duplicate?	Y	N	Y	Y	Y
6. Did the review authors perform data extraction in duplicate?	Y	N	Y	N	Y
7. Did the review authors provide a list of excluded studies and justify the exclusions?	Y	N	Y	N	N
8. Did the review authors describe the included studies in adequate detail?	PY	PY	PY	PY	PY
9. Did the review authors use a satisfactory technique for assessing the RoB in individual studies that were included in the review?					
RCT	PY	Y	Y	Y	Y
NRSI	NA	NA	NA	NA	NA
10. Did the review authors report on the sources of funding for the studies included in the review?	N	N	N	N	N
11. If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?					
RCT	Y	N	Y	Y	Y
NRSI	NA	NA	NA	NA	NA

12. If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?	Y	N	Y	N	N
13. Did the review authors account for RoB in individual studies when interpreting/discussing the results of the review?	Y	N	Y	N	N
14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?	Y	N	Y	Y	Y
15. If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?	Y	N	Y	Y	Y
16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?	Y	Y	Y	Y	Y
Overall rating	**Low***	**Critically low**	**Low***	**Critically low**	**Low**

Note. Y = yes, PY = partial yes, N = no. PICO: (P = population, I = intervention, C = comparator group, O = outcome). RCT = randomized controlled trial. NRSI = non-randomized studies of interventions. NA = not applicable. Items in bold are considered critical weaknesses if coded "no." All studies were triple coded and discrepancies were resolved by consensus to arrive at the final ratings.

3.2 Narrative Summary of AMSTAR2 Findings

The 5 included meta-analyses, taken together, are considered adequate for drawing reasonable conclusions about the combined efficacy of CM for SUD. The Bolívar et al. (2021) and Ginley et al. (2021) meta-analyses are considered the strongest with overall ratings of "low" based on fully- or partially-met criteria for all AMSTAR2 quality domains. However, these two meta-analyses, which included 76 of 85 (89%) total studies in the evaluation, were rated as low quality only because they did not justify restricting their systematic article search to English language publications only. This lack of justification may be overly penalizing because most empirical research on CM has been conducted in the United States (Miguel et al., 2016; Petry et al., 2017), and in all other respects, there were no other critical weaknesses in these two meta-analyses. Thus, it is more likely the overall ratings of these meta-analyses are "moderate."

According to the AMSTAR2 ratings of the other 3 meta-analyses, one was rated "low" quality (Sheridan Rains et al., 2020) because it did not account for risk of bias in individual studies when discussing results, and two were rated "critically low" quality (Destoop et al., 2021; Pfund et al., 2022) because they did not use a comprehensive literature search strategy, account for risk of bias in individual studies when discussing results, and provide a satisfactory explanation for, and discussion of, any heterogeneity in the results. See Table 2 for detailed AMSTAR2 ratings across all 5 meta-analyses.

Weighted summary effect size estimates were calculated using all 5 studies (regardless of the quality ratings from AMSTAR2), and sensitivity analyses were conducted with only the Bolívar et al. (2021) and Ginley et al. (2021) meta-analyses, excluding three of

the studies that were rated as “low” and “critically low” (Destoop et al., 2021; Pfund et al., 2022; Sheridan Rains et al., 2020).

4. Evaluating Outcomes and Judging Quality of the Evidence

4.1 Creating Procedures for Extracting Data

Two coders reviewed each meta-analysis and extracted effect sizes for every outcome, as well as their confidence intervals or standard errors, from the numerical tables accompanying forest plots. These outcomes reflected the effect of CM, relative to various control conditions (see Table 1).

The primary outcomes extracted were posttreatment and follow-up abstinence. Posttreatment reflected outcomes at treatment end or during treatment, and follow-up reflected outcomes at the longest available follow-up. Other outcomes included attendance, medication use, and cigarette smoking at treatment end/during treatment.

4.2 Data Collection and Validation

Data was entered directly into Comprehensive Meta-Analysis (CMA) version 3.3.070, and CMA was used to perform all effect size calculations. Cohen's *d* effect sizes (Cohen, 1988) quantified the effect of CM relative to control on all outcomes. Meta-analyses that reported effect sizes that were different from Cohen's *d*, like odds ratios (Ginley et al., 2021), were converted to Cohen's *d*.

4.3 Statistically Combining Effect Sizes

If a study was included in more than one meta-analysis, or if an individual study reported more than one effect size, the effect sizes were averaged into a single effect size (Borenstein et al., 2009; López-López et al., 2018). Across the 104 individual studies comprising the five meta-analyses, two studies were included in three different meta-analyses and 16 studies were included in two different meta-analyses. In total, 84 unique studies were included in the meta-analyses when accounting for duplicates.

4.4 Interpreting Results

A random effects analysis was used because considerable heterogeneity was expected in the effect sizes comprising each individual study. CM protocols involved different contingent targets (e.g., abstinence, attendance, employment, medication use) and reward schedules (i.e., different reward frequencies and magnitudes) which are known moderators of abstinence (Griffith et al., 2000; Lussier et al., 2006; Pfund et al., 2021).

Figure 2 depicts the forest plots of the effect sizes from the individual studies (k = 62, n = 9,199) comprising each meta-analysis for abstinence at posttreatment, and Figure 3 depicts the forest plot (k = 30, n = 4,614) for abstinence at follow up. The effect of CM relative to controls on posttreatment abstinence was d = 0.49, 95% CI [0.40, 0.59], and there was significant heterogeneity in effect sizes, $Q(61) = 204.93$, $p < .001$, $I^2 = 70.23$. The effect on follow-up abstinence was d = 0.09, 95% CI [0.01, 0.16], and there was no significant heterogeneity in effect sizes, $Q(30) = 39.51$, $p = 0.10$.

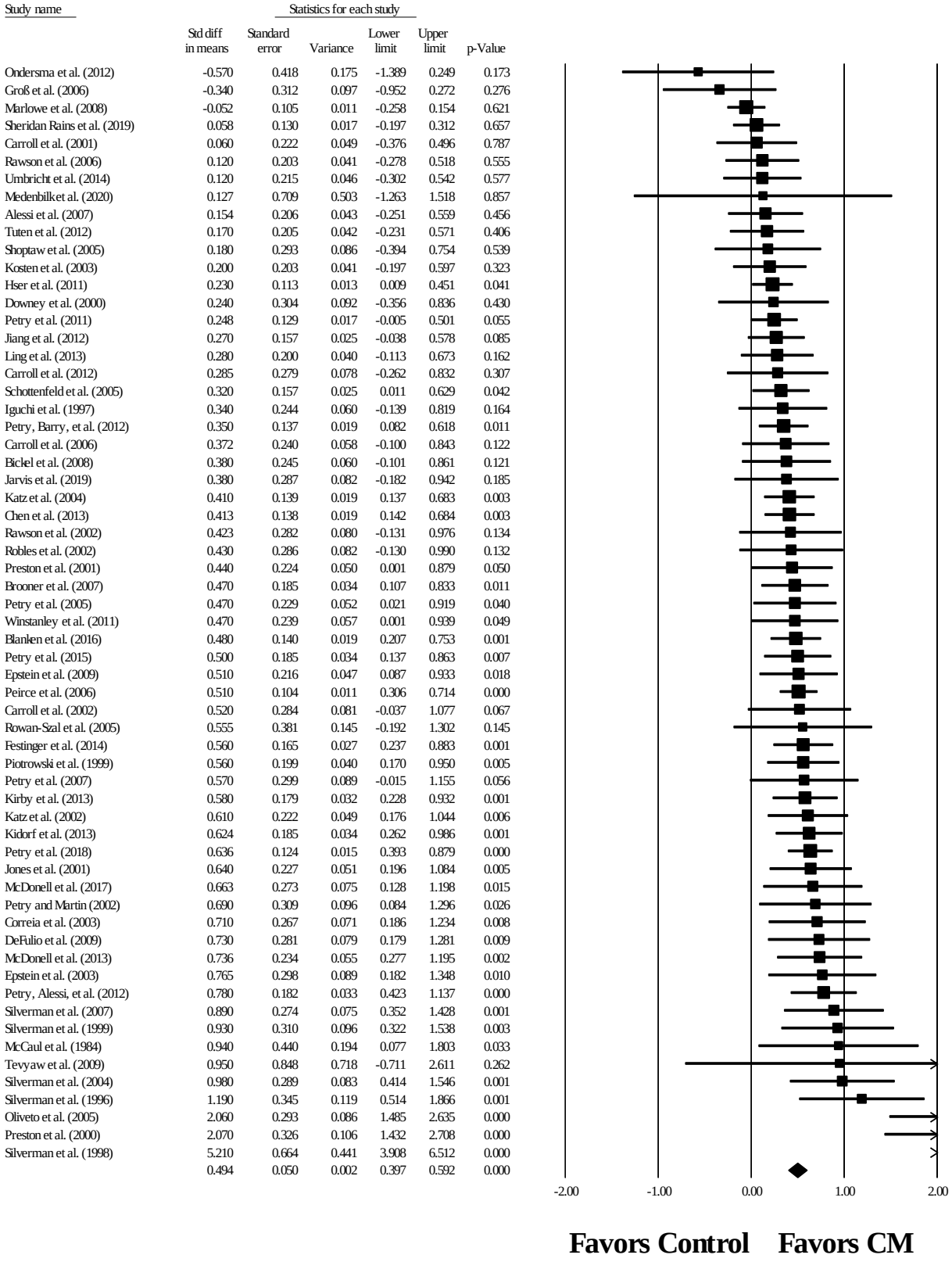

Figure 2. Forest Plot of Cohen's *d V*alues for Posttreatment Abstinence. This figure depicts Cohen's *d* values with 95% confidence intervals from the studies comprising meta-analyses that reported abstinence outcomes at treatment end or during treatment (Bolívar et al., 2021; Destoop et al., 2021; Pfund et al., 2022; Sheridan Rains et al., 2020).

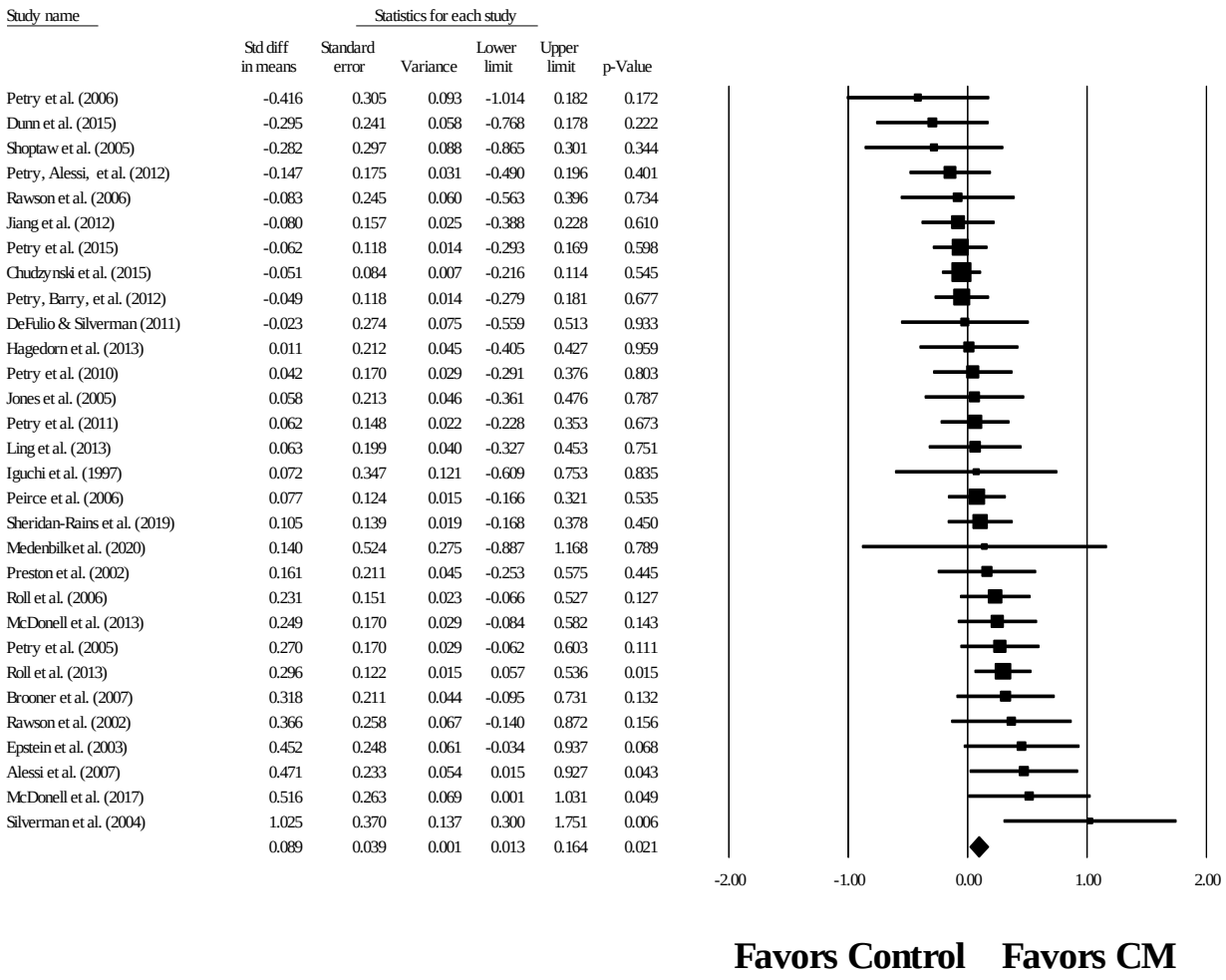

Figure 3. Forest Plot of Cohen's *d V*alues for Follow-Up Abstinence. This figure depicts Cohen's *d* values with 95% confidence intervals from studies comprising meta-analyses that reported abstinence outcomes at the longest available follow-up (Bolívar et al., 2021; Destoop et al., 2021; Ginley et al., 2021).

Figures 4, 5, and 6 depict the forest plots of the effect sizes from the individual studies comprising each meta-analysis for attendance (k = 17, n = 3,113), medication use (k = 9, n = 867), and cigarette smoking (k = 3, n = 219), respectively. The effect of CM relative to control on attendance was d = 0.44, 95% CI [0.27, 0.61], and there was significant heterogeneity in effect sizes, $Q(16) = 79.09$, $p < .001$, $I^2 = 79.77$. The effect of CM relative to control on medication use was d = 0.75, 95% CI [0.30, 1.21], and there was significant heterogeneity in effect sizes, $Q(8) = 60.93$, $p < .001$, $I^2 = 86.87$. The effect of CM relative to control on cigarette smoking was d = 0.79, 95% CI [0.43, 1.14], and there was no significant heterogeneity in effect sizes, $Q(2) = 2.71$, $p = 0.26$.

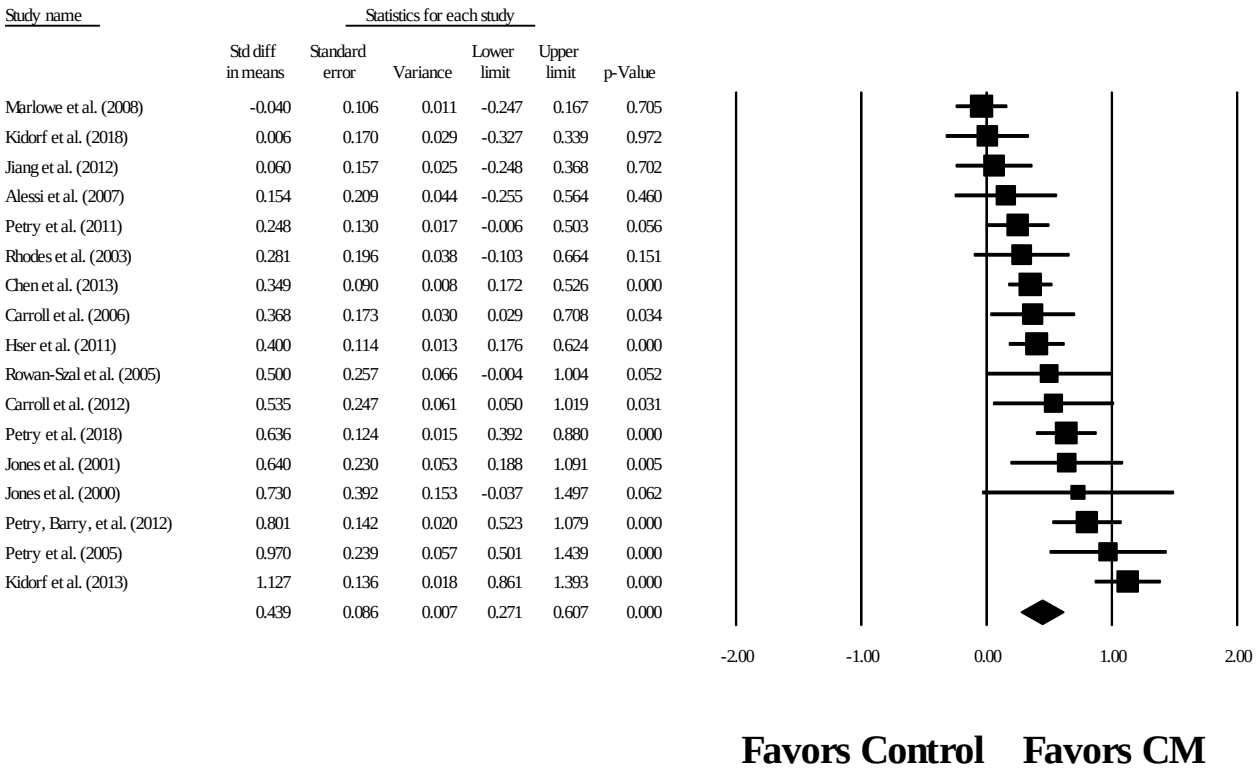

Figure 4. Forest Plot of Cohen's *d V*alues for Attendance. This figure depicts Cohen's *d* values with 95% confidence intervals from studies comprising meta-analyses that reported attendance outcomes at treatment end or during treatment (Bolívar et al., 2021; Pfund et al., 2022).

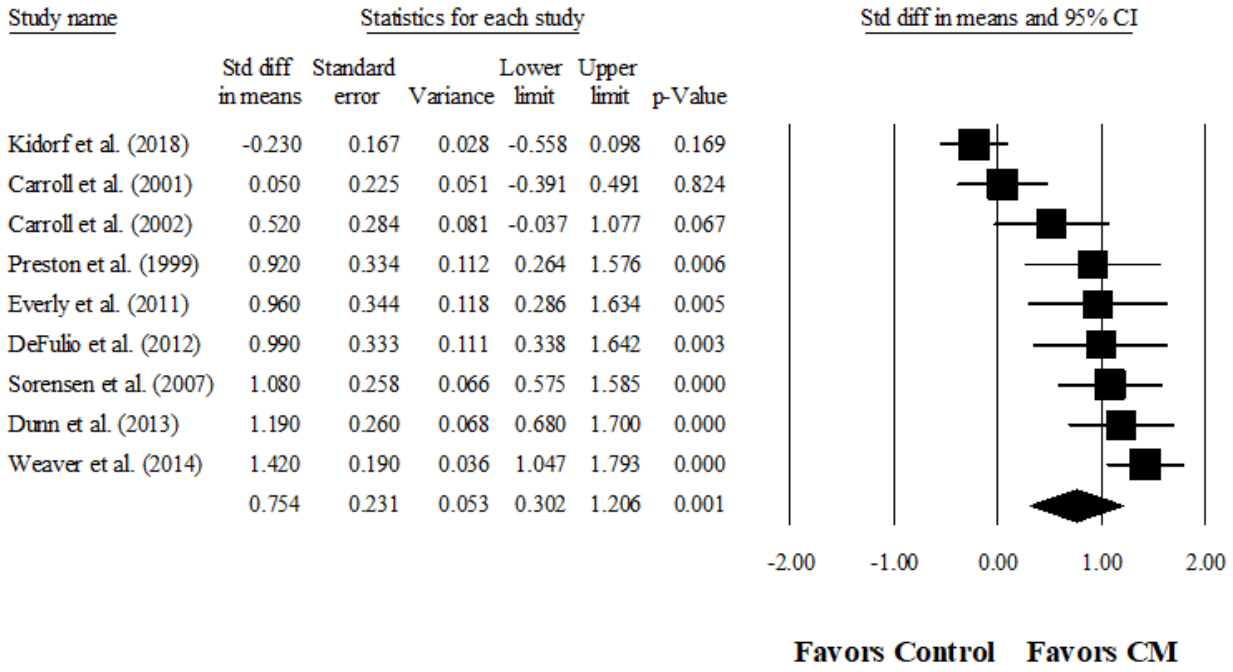

Figure 5. Forest Plot of Cohen's *d* Values for Medication Use. This figure depicts Cohen's *d* values with 95% confidence intervals from studies comprising meta-analyses that reported medication use outcomes at treatment end or during treatment (Bolívar et al., 2021).

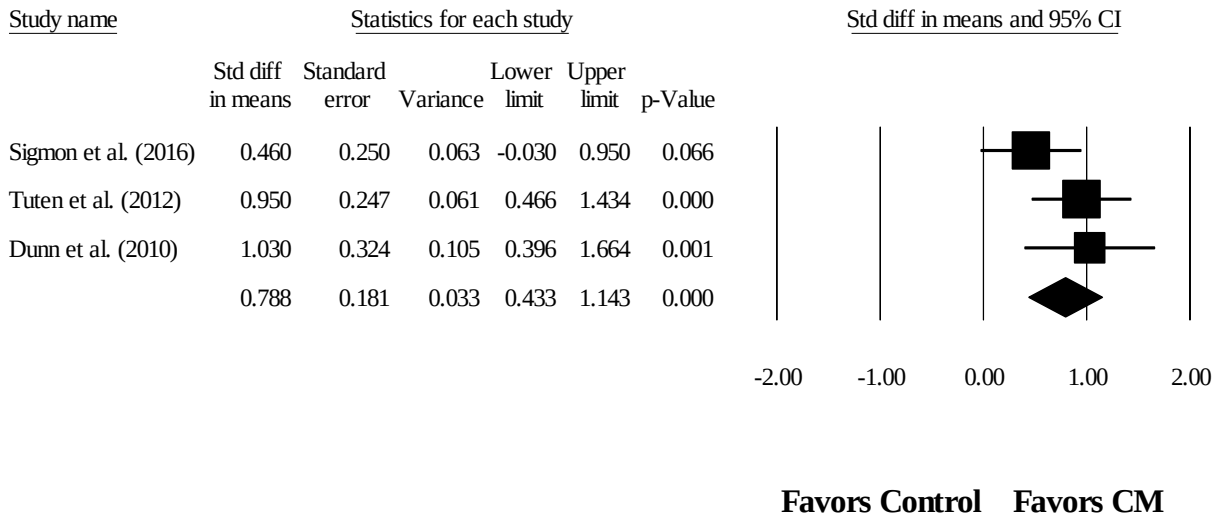

Figure 6. Forest Plot of Cohen's *d* Values for Cigarette Smoking. This figure depicts Cohen's *d* values with 95% confidence intervals from studies comprising meta-analyses that reported cigarette smoking outcomes at treatment end or during treatment (Bolívar et al., 2021).

Using Cohen's (1988) guidelines for classifying small effects as *d* = 0.2, medium effects as *d* = 0.5, and large effects as *d* = 0.8, the figures collectively illustrate that CM produces medium changes in abstinence at posttreatment and small changes in abstinence at follow-up. Regarding non-abstinence outcomes, the figures illustrate that CM produces medium changes in attendance and large changes in medication use and cigarette smoking. However, weighted effect sizes estimates for medication use and

cigarette smoking were comprised of comparatively fewer studies (9 and 3, respectively) than estimates for posttreatment abstinence (62), follow-up abstinence (30), and attendance (17). Thus, the effects of CM on medication use and cigarette smoking should be taken with caution.

4.5 Sensitivity Analyses

Several sensitivity analyses were conducted to examine the effect of CM, relative to control, on outcomes with three of the meta-analyses rated as “low” or “critically low” removed from the analyses (Figures 7, 8, and 9). The first analysis examined the effect of CM on posttreatment abstinence with the Destoop et al. (2021), Pfund et al. (2022), and Sheridan Rains et al. (2020) meta-analyses removed. The effect of CM relative to control on posttreatment abstinence (k = 51, n = 7,183) was d = 0.54, 95% CI [0.43, 0.64], and there was significant heterogeneity in effect sizes, $Q(50) = 161.18$, $p < .001$, $I^2 = 68.98$. The 0.54 effect size was slightly larger than the 0.49 effect size with the Destoop et al. (2021), Pfund et al. (2022), and Sheridan Rains et al. (2020) meta-analyses included.

The second analyses examine the effect of CM on follow-up abstinence with the Destoop et al. (2021) meta-analyses removed. The effect of CM relative to control on follow-up abstinence (k = 27, n = 4,147) was d = 0.08, 95% CI [0.00, 0.16]. There was not significant heterogeneity in effect sizes, $Q(26) = 35.59$, $p = 0.10$. The 0.08 effect size was almost equivalent to the 0.09 effect size with the Destoop et al. (2021) meta-analysis included.

The third analysis examined the effect of CM on attendance with the Pfund et al. (2022) meta-analysis removed. The effect of CM relative to control on attendance (k = 9, n = 1,357) was d = 0.45, 95% CI [0.22, 0.67], and there was significant heterogeneity in effect sizes, $Q(8) = 28.90$, $p < .001$, $I^2 = 72.32$. The 0.45 effect size was almost equivalent to the 0.044 effect size with the Pfund et al. (2022) meta-analysis included. Sensitivity analyses were not conducted on medication use, cigarette smoking, or follow-up abstinence because no meta-analyses rated “low” or “critically low” reported these outcomes.

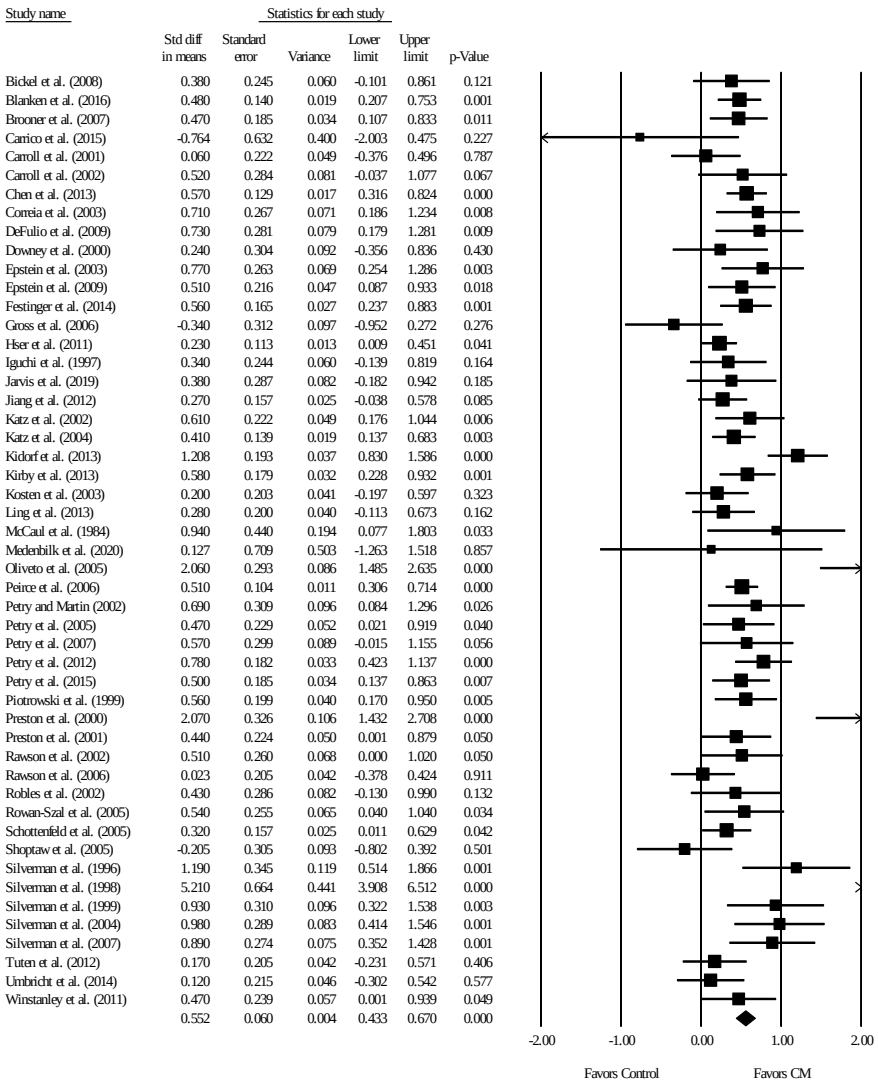

Figure 7. Forest Plot of Cohen's *d V*alues for Posttreatment Abstinence. This figure depicts Cohen's *d* values with 95% confidence intervals from studies comprising "moderate" quality meta-analyses that reported abstinence outcomes at treatment end or during treatment (Bolívar et al., 2021).

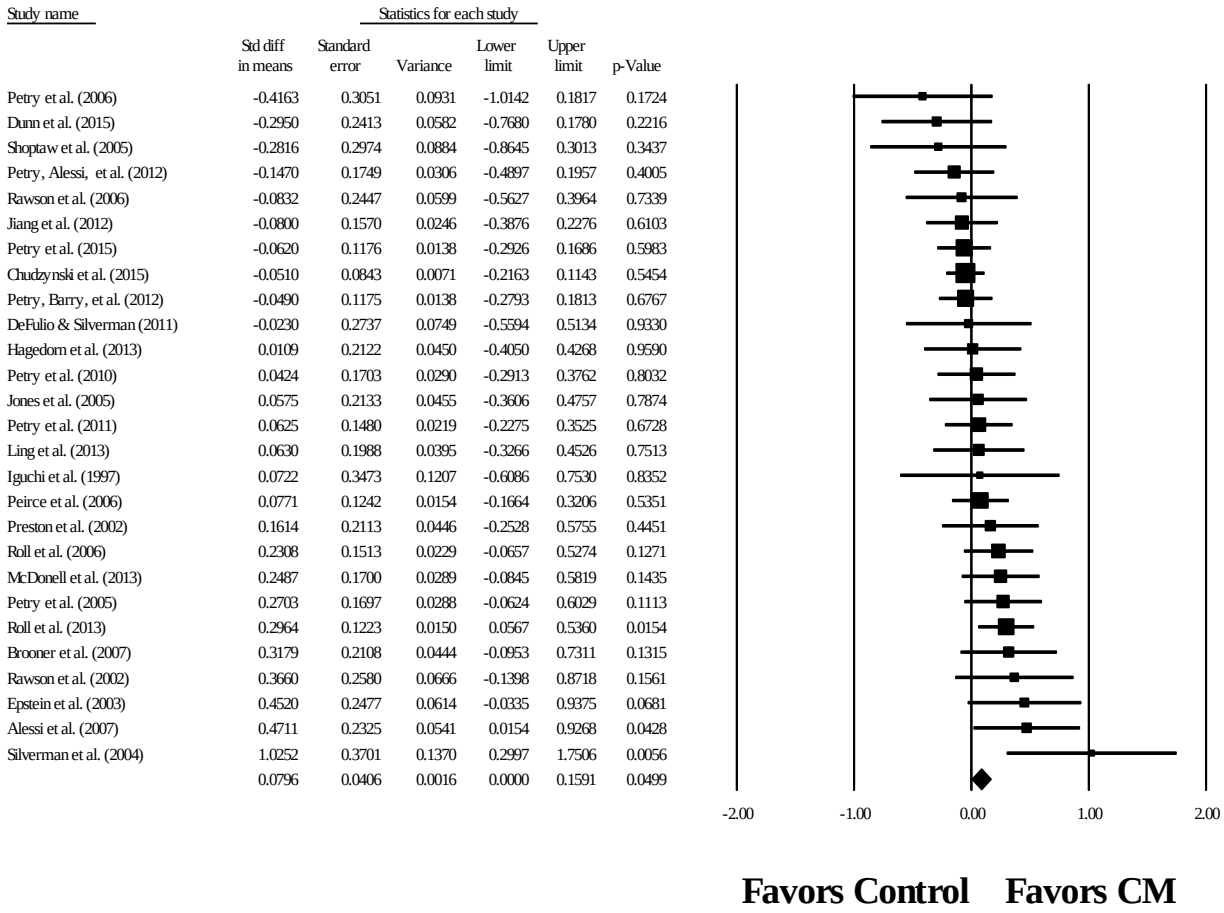

Figure 8. Forest Plot of Cohen's *d V*alues for Follow-Up Abstinence. This figure depicts Cohen's *d* values with 95% confidence intervals from studies comprising "moderate" quality meta-analyses that reported abstinence outcomes at treatment end or during treatment (Bolívar et al., 2021; Ginley et al., 2021).

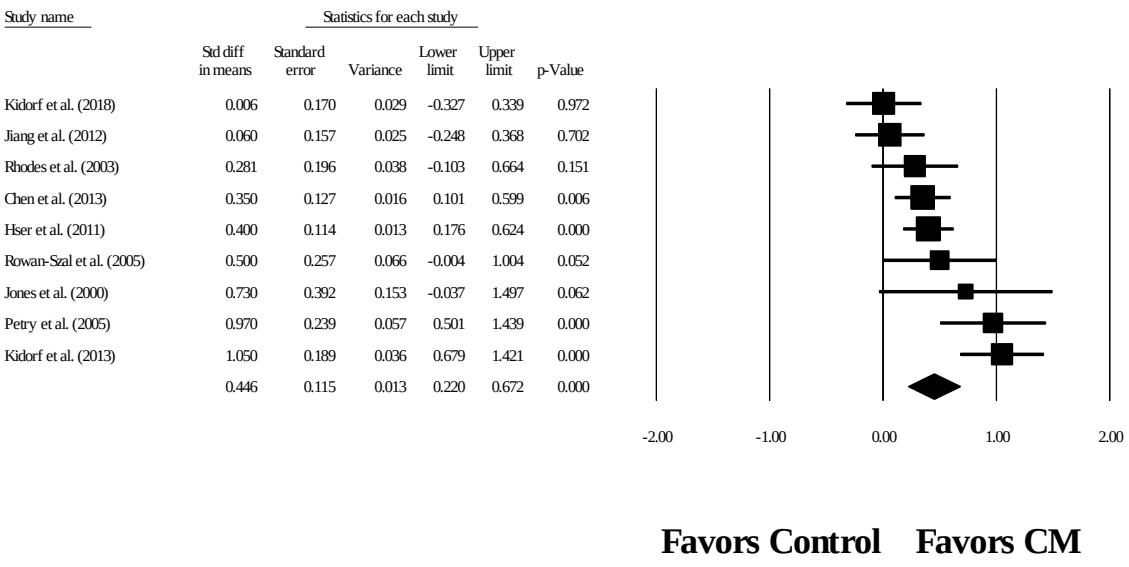

Figure 9. Forest Plot of Cohen's *d V*alues for Attendance. This figure depicts Cohen's *d* values with 95% confidence intervals from studies comprising "moderate" quality meta-analyses that reported attendance outcomes at treatment end or during treatment (Bolívar et al., 2021).

4.6 Judging the Quality of the Evidence

Table 3

Judging the Quality of the Evidence for CM for SUD

Quality	Criteria
✓ *High quality*	All of the following: • There is a wide range of studies included in the analyses with no major limitations. • There is little variation between studies. • The summary estimate has a narrow confidence interval.
☐ *Moderate quality*	At least one of the following: • There are only a few studies, and some have limitations but not major flaws. • There is some variation between studies, or the confidence interval of the summary estimate is wide.
☐ *Low quality*	Any of the following: • The studies have major flaws. • There is important variation between studies. • The confidence interval of the summary estimate is very wide.

Note. We are confident that there are no major limitations in either the Bolívar et al. (2021) or Ginley et al. (2021) meta-analyses, and these meta-analyses included 74 of 84 (88%) total studies in our calculation of weighted effect size estimates. Although there was significant heterogeneity in the effect sizes of individual studies from these two meta-analyses ($Q(50) = 161.18$, $p < .001$, $I^2 = 68.98$, this heterogeneity was expected because the CM protocols involved different contingent targets (i.e., abstinence, attendance, employment, medication use) and reward schedules (i.e., different reward frequencies and magnitudes), which are known moderators of posttreatment abstinence in CM (Griffith et al., 2000; Lussier et al., 2006; Pfund et al., 2022). There was little variation in the effect sizes for follow-up abstinence ($Q(31) = 38.96$, $p = .13$). Furthermore, the weighted summary effect size estimates at posttreatment and follow-up had narrow confidence intervals (95% CI [0.43, 0.64] and 95% CI [0.01, 0.16], respectively).

5. Consideration of Additional Contextual Factors

Table 4

Additional Contextual Factors Considered in Increasing or Decreasing the GRADE Recommendation for CM for SUD

Positive	Negative
✓ Treatment appears superior to other established and effective treatment(s)	☐ There are other psychological treatments that have well-documented and much larger effects
✓ The treatment generates an effect that is similar to other well-studied treatments, but requires a very small number of sessions or length of time to generate the same effect at a much lower cost	☐ The treatment generates an effect that is similar to other well-studied treatments, but requires a very large number of sessions or length of time to generate the same effect at a much higher cost
☐ Evidence supports the purported mechanism or active ingredient(s) of treatment	☐ Evidence fails to support the purported mechanism or active ingredient(s) of treatment
✓ Treatment has demonstrated good effects with marginalized groups	☐ Treatment has demonstrated weak effects with marginalized groups
☐ Treatment has been studied by a wide array of researchers without strong allegiance to the treatment	☐ Treatment has been studied by a narrow array of researchers with strong allegiance to the treatment
☐ Other:	☐ Other:

Note. This table identifies additional positive contextual factors supported by the literature on CM for SUD and was adapted from Tolin et al. (2015). Lack of identification of a positive or negative assessment of a contextual factor indicates that there are not enough data to make a firm conclusion in this category for CM.

5.1 Consideration of Contextual Factors

Several contextual factors were considered before providing an overall recommendation on CM as an empirically supported treatment for SUD. The effect of CM on posttreatment abstinence is larger than other established treatments (i.e., various forms of pharmacotherapy and psychotherapy), especially for cocaine use disorder treatment and when CM is combined with cognitive-behavioral therapy (Bentzley et al., 2021; Dutra et al., 2008). The effect of CM on follow-up abstinence is equal to or larger than the effect of other well-studied treatments, like cognitive-behavioral therapy (Ginley et al., 2021; Magill et al., 2019; Magill & Ray, 2009).

CM also demonstrates good effects with historically marginalized groups, including African American/Black, American Indian, Alaskan Native, and Hispanic people (Hirchak et al., 2021; Pfund et al., 2022; Venner et al., 2021), as well as lesbian, gay, and bisexual people (Zajac et al., 2020). Secondary analyses of clinical trials indicate comparable efficacy in CM among African American, Hispanic, and White people (Barry et al., 2009).

There is currently limited evidence to support the purported mechanism of CM. The purported mechanism is the contingent-incentive system (Witkiewitz et al., 2022), and some studies support that increases in abstinence during treatment mediates the association between the contingent-incentive system and outcomes (i.e., quality of life, psychiatric symptoms) in the months after CM termination (Petry et al., 2007, 2013). However, to our knowledge, no mechanistic studies have been conducted on the

contingent-incentive system for non-abstinent target behaviors (e.g., treatment attendance or medication use).

There is also limited evidence to determine that CM has been studied by a wide array of researchers without strong allegiance to CM. Over 10 different corresponding authors who represented unique universities were listed across the 85 unique studies in the present evaluation. Authors sometimes indicated that they used CM protocols that they created, but no studies have been conducted to assess whether this allegiance or other dimensions of allegiance (e.g., reprint analysis, examinations of previous publications on CM, interviews with colleagues of authors, or interviews with the authors themselves) was related to outcomes.

6. Overall Treatment Recommendation

Table 5

Overall Treatment Recommendation for CM for SUD

Recommendation	Criteria
☐ *Very strong recommendation*	All of the following: • There is high-quality evidence that the treatment produces a clinically meaningful effect on symptoms of the disorder being treated • There is high-quality evidence that the treatment produces a clinically meaningful effect on functional outcomes • There is high-quality evidence that the treatment produces a clinically meaningful effect on symptoms and/or functional outcomes at least three months after treatment discontinuation • At least one well-conducted study has demonstrated effectiveness in non-research settings
✓ *Strong recommendation*	At least one of the following: • There is moderate- to high-quality evidence that the treatment produces a clinically meaningful effect on symptoms of the disorder being treated • There is moderate- to high-quality evidence that the treatment produces a clinically meaningful effect on functional outcomes
☐ *Weak recommendation*	Any of the following: • There is only low- or very low-quality evidence that the treatment produces a clinically meaningful effect on symptoms of the disorder being treated • There is only low- or very low-quality evidence that the treatment produces a clinically meaningful effect on as well as on functional outcomes • There is moderate- to high-quality evidence that the effect of the treatment, although statistically significant, may not be of a magnitude that is clinically meaningful

Note. This table was adapted from Tolin et al. (2015).

6.1 Overall Treatment Recommendation

There is high-quality evidence that CM for SUD produces a clinically meaningful and statistically significant effect on posttreatment and follow-up abstinence, as evidenced by the Bolivar et al. (2021) and Ginley et al. (2021) meta-analyses. As such, based on the criteria outlined by Tolin and colleagues (2015), the current status of the literature merits a **strong** recommendation for CM in the treatment of SUD.

Several contextual factors further strengthen our **strong** recommendation:

1. CM generates a strong effect on posttreatment abstinence that is greater than the effects of other treatments (Bentzley et al., 2021).
2. CM generates a small effect on follow-up abstinence that is similar or better than the effects of other treatments like CBT (Ginley et al., 2021; Magill et al., 2019; Magill & Ray, 2009).

3. CM has demonstrated good effects with historically marginalized groups (e.g., individuals identifying as African American/Black, American Indian, Alaskan Native, and Hispanic) (Hirchak et al., 2021; Pfund et al., 2022; Venner et al., 2021).

We cannot provide a “very strong recommendation” for CM as an empirically supported treatment for SUD for two reasons. First, there is limited understanding about the effect of CM on functional outcomes (e.g., psychosocial functioning) at posttreatment and after the discontinuation of incentives. Second, little is known about the effectiveness of CM in real-world settings. Future research should examine the effect of CM on outcomes, such as psychiatric symptoms and quality of life. Future research should also examine how CM is implemented in real-world settings and how treatment providers in these settings conduct CM in ways that are consistent with how CM is conducted in research settings.

References

Barry, D., Sullivan, B., & Petry, N. M. (2009). Comparable efficacy of contingency management for cocaine dependence among African American, Hispanic, and White methadone maintenance clients. *Psychology of Addictive Behaviors*, *23*, 168–174. https://doi.org/10.1037/a0014575.Comparable

Bentzley, B. S., Han, S. S., Neuner, S., Humphreys, K., Kampman, K. M., & Halpern, C. H. (2021). Comparison of treatments for cocaine use disorder among adults: A systematic review and meta-analysis. *JAMA Network Open*, *4*, 1–22. https://doi.org/10.1001/jamanetworkopen.2021.8049

Bolívar, H. A., Klemperer, E. M., Coleman, S. R. M., DeSarno, M., Skelly, J. M., & Higgins, S. T. (2021). Contingency management for patients receiving medication for opioid use disorder. *JAMA Psychiatry*, *78*, 1092–1102. https://doi.org/10.1001/jamapsychiatry.2021.1969

Boness, C. L., Hershenberg, R., Grasso, D. J., Kaye, J., Mackintosh, M.-A., Nason, E., Shah, A., & Raffa, S. D. (2021). *The Society of Clinical Psychology's manual for the evaluation of psychological treatments using the Tolin Criteria*. https://doi.org/doi.org/10.31219/osf.io/8hcsz

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. John Wiley & Sons.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Routledge.

Destoop, M., Docx, L., Morrens, M., & Dom, G. (2021). Meta-analysis on the effect of contingency management for patients with both psychotic disorders and substance use disorders. *Journal of Clinical Medicine*, *10*. https://doi.org/10.3390/jcm10040616

Dutra, L., Stathopoulou, G., Basden, S. L., Leyro, T. M., Powers, M. B., & Otto, M. W. (2008). A meta-analytic review of psychosocial interventions for substance use disorder. *American Journal of Psychiatry*, *165*, 179–187. https://doi.org/10.1016/S0168-9002(99)01439-4

Ginley, M. K., Pfund, R. A., Rash, C. J., & Zajac, K. (2021). Long-term efficacy of contingency management treatment based on objective indicators of abstinence from illicit substance use up to 1 year following treatment: A meta-analysis. *Journal of Consulting and Clinical Psychology*, *89*(1), 58–71. https://doi.org/10.1037/ccp0000552

Griffith, J. D., Rowan-Szal, G. A., Roark, R. R., & Simpson, D. D. (2000). Contingency management in outpatient methadone treatment: a meta-analysis. *Drug and Alcohol Dependence*, *58*(1–2), 55–66. https://doi.org/10.1016/s0376-8716(99)00068-x

Hirchak, K. A., Lyons, A. J., Herron, J. L., Kordas, G., Shaw, J. L., Jansen, K., Avey, J. P., McPherson, S. M., Donovan, D., Roll, J., Buchwald, D., Ries, R., & McDonell, M. G. (2021). Contingency management for alcohol use disorder reduces cannabis use among American Indian and Alaska Native adults. *Journal of Substance Abuse Treatment*, 108693. https://doi.org/10.1016/j.jsat.2021.108693

López-López, J. A., Page, M. J., Lipsey, M. W., & Higgins, J. P. T. (2018). Dealing with effect size multiplicity in systematic reviews and meta-analyses. *Research Synthesis Methods*, *9*, 336–351. https://doi.org/10.1002/jrsm.1310

Lussier, J. P., Heil, S. H., Mongeon, J. A., Badger, G. J., & Higgins, S. T. (2006). A meta-analysis of voucher-based reinforcement therapy for substance use disorders. *Addiction*, *101*, 192–203. https://doi.org/10.1111/j.1360-0443.2006.01311.x

Magill, M., & Ray, L. A. (2009). Cognitive-behavioral treatment with adult alcohol and illicit drug users: A meta-analysis of randomized controlled trials. *Journal of Studies on Alcohol and Drugs*, *70*, 516–527. https://doi.org/10.15288/jsad.2009.70.516

Magill, M., Ray, L., Kiluk, B., Hoadley, A., Bernstein, M., Scott Tonigan, J., & Carroll, K. (2019). A meta-analysis of cognitive-behavioral therapy for alcohol or other drug use disorders: Treatment efficacy by contrast condition. *Journal of Consulting and Clinical Psychology*, *87*, 1093–1105. https://doi.org/10.1037/ccp0000447

Miguel, A. Q. C., Madruga, C. S., Cogo-Moreira, H., Yamauchi, R., Simões, V., da Silva, C. J., McPherson, S., Roll, J. M., & Laranjeira, R. R. (2016). Contingency management is effective in promoting abstinence and retention in treatment among crack cocaine users with a previous history of poor treatment response: a crossover trial. *Psychology of Addictive Behaviors*, *30*,

536–543. https://doi.org/10.1186/s41155-019-0127-2
Montgomery, L. T., Carroll, K. M., & Petry, N. M. (2015). Initial abstinence status and contingency management treatment outcomes: Does race matter? *Journal of Consulting and Clinical Psychology*, *83*, 473–481. https://doi.org/10.1037/a0039021
Petry, N. M., Alessi, S. M., & Hanson, T. (2007). Contingency management improves abstinence and quality of life in cocaine abusers. *Journal of Consulting and Clinical Psychology*, *75*, 307–315. https://doi.org/10.1037/0022-006X.75.2.307
Petry, N. M., Alessi, S. M., Olmstead, T. A., Rash, C. J., & Zajac, K. (2017). Contingency management treatment for substance use disorders: How far has it come, and where does it need to go? *Psychology of Addictive Behaviors*, *31*, 897–906. https://doi.org/10.1037/adb0000287
Petry, N. M., Alessi, S. M., & Rash, C. J. (2013). Contingency management treatments decrease psychiatric symptoms. *Journal of Consulting and Clinical Psychology*, *81*, 926–931. https://doi.org/10.1037/a0032499
Pfund, R. A., Ginley, M. K., Rash, C. J., & Zajac, K. (2022). Contingency management for treatment attendance: A meta-analysis. *Journal of Substance Abuse Treatment*, *133*, 108556. https://doi.org/10.1016/j.jsat.2021.108556
Schardt, C., Adams, M. B., Owens, T., Keitz, S., & Fontelo, P. (2007). Utilization of the PICO framework to improve searching PubMed for clinical questions. *BMC Medical Informatics and Decision Making*, *7*, 1–6. https://doi.org/10.1186/1472-6947-7-16
Shea, B. J., Reeves, B. C., Wells, G., Thuku, M., Hamel, C., Moran, J., Moher, D., Tugwell, P., Welch, V., Kristjansson, E., & Henry, D. A. (2017). AMSTAR 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*, *358*, 1–9. https://doi.org/10.1136/bmj.j4008
Sheridan Rains, L., Steare, T., Mason, O., Johnson, S., Rains, L. S., Steare, T., Mason, O., & Johnson, S. (2020). Improving substance misuse outcomes in contingency management treatment with adjunctive formal psychotherapy: a systematic review and meta-analysis. *BMJ Open*, *10*(10), e034735. https://doi.org/10.1136/bmjopen-2019-034735
Tolin, D. F., McKay, D., Forman, E. M., Klonsky, E. D., & Thombs, B. D. (2015). Empirically supported treatment: Recommendations for a new model. *Clinical Psychology: Science and Practice*, *22*, 317–338. https://doi.org/10.1111/cpsp.12122
Venner, K. L., Serier, K., Sarafin, R., Greenfield, B. L., Hirchak, K., Smith, J. E., & Witkiewitz, K. (2021). Culturally tailored evidence-based substance use disorder treatments are efficacious with an American Indian Southwest tribe: An open-label pilot-feasibility randomized controlled trial. *Addiction*, *116*, 949–960. https://doi.org/10.1111/add.15191
Witkiewitz, K., Pfund, R. A., & Tucker, J. A. (2022). Mechanisms of behavior change in substance use disorder with and without formal treatment. *Annual Review of Clinical Psychology*, *18*, 1–29. https://doi.org/10.1146/annurev-clinpsy-072720-014802
Zajac, K., Rash, C. J., Ginley, M. K., & Heck, N. C. (2020). Sexual orientation and substance use treatment outcomes across five clinical trials of contingency management. *Psychology of Addictive Behaviors : Journal of the Society of Psychologists in Addictive Behaviors*, *34*(1), 128–135. https://doi.org/10.1037/adb0000494