Psicologia: Teoria e Pesquisa
2022, v.38, e38215

DOI: https://doi.org/10.1590/0102.3772e38215.en

Behavioral Sciences

Contingency Management applied to Alcohol Use Disorder: Systematic Review*

Karina de Souza Silva[1,**], Angelo Augusto Silva Sampaio[1], & André de Queiroz Constantino Miguel[2]

[1]*Universidade Federal do Vale do São Francisco, Petrolina, PE, Brasil*

[2]*Universidade Federal de São Paulo, São Paulo, SP, Brasil*

ABSTRACT – This systematic review evaluated the efficacy of applying Contingency Management (CM) to Alcohol Use Disorder. We followed the PRISMA recommendation and consulted the following databases: Cochrane Library, MEDLINE Complete, PsycINFO and Pubmed. A total of eight randomized controlled trials were included in this review, all of them with good methodological quality. In seven of these, CM was more efficacious in promoting continuous abstinence. Both trials that evaluated treatment retention found statistically significant results favorable to CM. On two of the three trials presenting follow-up results, CM was more efficacious in promoting abstinence. The large-scale application of CM can promote substantial public health improvements and should be encouraged.

KEYWORDS: contingency management, alcohol use disorder, evidence-based practice, systematic review, randomized controlled trials

Manejo de Contingência aplicado ao Transtorno por Uso de Álcool: Revisão Sistemática

RESUMO – Esta revisão sistemática avaliou a eficácia do Manejo de Contingência (MC) no tratamento do Transtorno por Uso de Álcool. Para isso, foi utilizada a recomendação PRISMA e consultadas as bases de dados: Cochrane Library, MEDLINE Complete, PsycINFO e Pubmed. Foram incluídos oito ensaios clínicos randomizados nesta revisão. Em sete, o MC foi mais eficaz em promover abstinência continuada. Dos dois que avaliaram a retenção no tratamento, ambos encontraram resultados estatisticamente favoráveis ao MC. Dos três que apresentaram resultados de avaliação de seguimento, em dois o MC foi mais eficaz em promover abstinência. Todos apresentaram boa qualidade metodológica. A aplicação do MC em larga escala pode promover melhorias substanciais para a saúde pública e deve ser encorajada.

PALAVRAS-CHAVE: manejo de contingência, transtorno por uso de álcool, práticas baseadas em evidências, revisão sistemática, ensaios clínicos randomizados

Alcohol use is related to several relevant health and social problems. In Brazil, the prevalence of alcohol use disorder (AUD) has increased in recent decades and it is estimated that 4.2% of Brazilians (6.9% of men and 1.6% of women) meet criteria for this diagnosis (World Health Organization [WHO], 2018). Alcohol consumption in the country is associated with fatal car accidents (de Carvalho Ponce et al., 2011), violence (Abdalla et al., 2018), rape (Massaro et al., 2019), homicide (Andreuccetti et al., 2018; 2009) and suicide (Ponce et al., 2008). Within a year, more than 72,000 deaths (5.5% of all deaths) were attributable to alcohol consumption, making it the factor that most contributed to the country's mortality and morbidity burden (Degenhardt et al., 2018; WHO, 2018).

* Article derived from the Course Completion Work defended by the first author and supervised by the two other authors, in the undergraduate course in Psychology at the Federal University of Vale do São Francisco.

** Email: souza.kaah12@gmail.com

Submetido: 10/06/2020; Aceito: 09/06/2021.

According to the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), AUD consists of the continued use of alcohol despite significant substance-related problems (American Psychiatric Association [APA], 2014). Diagnosis is defined by the presence, for at least one year, of two or more (out of 11) behavioral or physiological symptoms related to: decreased control over substance use (e.g., craving, too much time spent in use-related activities), social impairment (e.g., activities abandonment or reduction, work problems), risky use (e.g., danger to physical integrity) and pharmacological criteria (tolerance and abstinence). AUD is also commonly accompanied by comorbidities such as depression, anxiety, insomnia, conduct problems and suicide risk (APA, 2014).

Given the high prevalence and severity of AUD, pharmacological (e.g., disulfiram, naltrexone) and psychosocial interventions have been developed for its treatment. Among the latter, cognitive and behavioral treatments present the best evidence of efficacy (Diehl et al., 2019; McCrady, 2016). In a recent meta-analysis of randomized controlled trials (RCTs) comparing the effects of various psychosocial interventions in promoting alcohol abstinence among individuals with AUD, Gao et al. (2018) observed that Contingency Management (CM) had the best efficacy measures during treatment.

CM is a behavioral intervention based on the principle of operant reinforcement. Basically, it consists of changing the user's current environment in order to increase the presence of reinforcers contingent on responses that are alternative or incompatible to substance use (Petry, 2011). It involves systematically and objectively monitoring substance use, preferably using biological markers (e.g., through a breathalyzer), and immediately and abundantly reinforcing the production of evidence that there was no consumption – therefore, that the individual was emitting abstinence-related responses, necessarily incompatible with substance consumption. Abstinence is the main target behavior in most studies—given the ease of objectively assessing it—however it is also common for reinforcement to be presented contingent on adherence to pharmacological treatment, participation in treatment-related activities, among others (Petry, 2011).

Recently, Miguel et al. (2016, 2017, 2018, 2019) carried out the first RCT to assess CM efficacy in Brazil—for the treatment of crack cocaine use disorder, and also including CM intervention in alcohol consumption. The intervention lasted 12 weeks and was carried out in a Medical Specialties Outpatient Clinic (*Ambulatório Médico de Especialidades*, AME) at São Paulo (SP) with 65 crack cocaine users in a high social vulnerability situation: most of them were unemployed (83.1%) and had already presented risky behavior, such as sleeping on the streets (64.6%) and attending "Crackland" (89.2%). The control group was offered the institution's standard treatment and the experimental group, standard treatment plus CM. All participants were encouraged to leave urine and breathalyzer samples three times a week. In the experimental group, when the urine test for cocaine and crack cocaine was negative, the participant received vouchers with a certain monetary value for being abstinent. These vouchers were exchanged for products available in the community within a 1 km radius from the service, accompanied by one of the researchers. If the participant was abstinent from crack cocaine and tested negative on the breathalyzer, he received extra vouchers. The main results showed an increase in treatment adherence and participation, reduction in crack cocaine use, and promotion of crack cocaine continuous abstinence among participants in the experimental group compared to those who received only the standard treatment (Miguel et al., 2016, 2019). Secondary outcomes of lower alcohol consumption were also observed among participants who received CM, in addition to a reduction in symptoms of depression and anxiety (Miguel et al., 2016, 2017).

In recent years, several meta-analyses and literature reviews have consistently pointed to CM efficacy and effectiveness in the treatment of substance use disorders (SUD), particularly among stimulant substances (e.g., cocaine, crack, methamphetamine; Davis et al., 2016; Lussier et al., 2006; Prendesgast et al., 2006). As a result of this substantial evidence of efficacy, in 1998 the US National Institute on Drug Abuse (NIDA) published a therapeutic manual detailing how to implement CM in open cocaine treatment services (NIDA, 1998). In 2007, the United Kingdom's National Institute for Health and Clinical Excellence (NICE) recommended the inclusion of CM in the country's National Substance Abuse Treatment Agency (Pilling et al., 2007). Despite the effectiveness of CM in the treatment of stimulant substances, few studies involving this intervention have been carried out exclusively to assess its application to AUD. In a non-systematic review of the literature, Wong et al. (2008) pointed to the likely effectiveness of the CM for this audience, but emphasized that the studies carried out so far had serious methodological problems. The meta-analysis by Gao et al. (2018), on the other hand, focused on comparing different interventions for SUD. In order to overcome this gap, the present study aimed to compile evidence of CM efficacy in AUD treatment through a systematic review of the literature on RCTs involving CM applied to the adult population with this diagnosis.

METHOD

Database Search and Selection Strategies

This systematic review was elaborated following the Preferred Reporting Items recommendation for Systematic Reviews and Meta-Analyses (PRISMA; Moher et al., 2009), which consists of a checklist with 27 items that should compose the review and a four-step flow diagram model (identification, selection, eligibility and inclusion), aiming to improve reporting and reduce publication bias (Moher et al., 2009).

Due to the pyramid of evidence used on the medical sciences, where RCT design is considered the "gold standard" to assess the efficacy of one or more interventions (Kendall, 2003), only RCTs involving CM for AUD were considered in this review. An RCT consists of selecting a sample of participants with the same diagnosis and randomly distributing it between an experimental group, which receives the intervention to be evaluated, and a control (or comparison) group, which receives an alternative (conventional) treatment (Kendall, 2003). To find the RCTs, we consulted Cochrane Library, MEDLINE Complete, PsycINFO and Pubmed databases, which were chosen based on the area literature. Combinations of the terms "alcohol" OR "drinker" OR "drinking" AND "contingency management", which should necessarily appear in the title and/or abstract, were used in the searches. To facilitate study identification, the option "clinical trials" was selected in all databases, considering only studies that were part of this category. All studies published between January 2000 and April 2020 were included.

After identifying the studies in the databases, we excluded: duplicates, unpublished material (e.g., research projects, conference presentations) and studies that did not present isolated CM as an independent variable (i.e., studies in which CM was implemented along with other interventions, such as Community Reinforcement Therapy). This initial inclusion and exclusion process was carried out by analyzing the RCTs titles and abstracts. The studies Method section was consulted in case of any dubiety. RCTs where alcohol consumption was not part of the evaluated outcomes and/or CM reinforcement was also dependent on other substances abstinence were also excluded. In addition, RCTs in which the control group was an CM adaptation were excluded.

Studies Characterization

Selected articles were categorized according to: authorship, year of publication, country in which the research was conducted, study setting, recruitment, aspects related to the sample (size, population, diagnosis and comorbidities), biological markers collection frequency, control-group type of treatment, incentives types, reinforcement schedule, treatment duration and follow-up duration (if applicable).

CM Efficacy Analysis

Three CM efficacy outcomes were analyzed: abstinence, binge drinking (i.e., consumption of at least four doses on a single occasion for women and five doses for men) and treatment retention (i.e., consecutive days/weeks of service attendance). Due to the heterogeneity of instruments (i.e., breathalyzer, urinalysis, transdermal sensors and self-report) and measurements (i.e., percentage of negative tests submitted, percentage of abstinence days, number of days until a lapse and longest continuous abstinence sequence) for assessing abstinence, this outcome was subdivided in:

1. *Percentage of abstinence days,* which included: (a) total frequency/percentage of negative breathalyzer testing; (b) total frequency of negative urine tests; and (c) percentage of abstinence days according to self-report; and
2. *Longest continuous abstinence sequence*, defined as the longest sequence of continuous abstinence measured through consecutive submission of biological markers. We also included in this outcome the time interval between the beginning of treatment and a lapse.

CM efficacy was evaluated separately for results during treatment and follow-up.

Methodological Quality Assessment

The RCTs methodological quality assessment was carried out using the Cochrane Risk of Bias tool (Higgins et al., 2011). This instrument assesses RCTs risk of bias in the following domains: (1) randomization process; (2) deviations from the intended interventions; (3) missing outcome data; (4) outcome measurement; and (5) selection of reported results. Each domain is evaluated individually using algorithms. The risk of general bias is determined based on the judgment of these domains, being classified into:

- *low risk of bias*: when all domains had minimal risk;
- *some concerns*: when at least one domain presented this classification, but none presented high risk; or
- *high risk of bias*: when at least one out of the five domains presented high risk of bias or some risk for multiple domains.

RESULTS

Figure 1 shows the selection process steps for the RCTs included in this review. We identified 615 articles in the databases. After the clinical trials option was selected, 368 were excluded. Next, 74 duplicate articles were excluded and, based on the reading of titles and abstracts, another 111 articles that did not meet selection criteria. Eight out of the 62 articles selected for eligibility assessment met the inclusion criteria.

Studies Characterization

Table 1 presents the RCTs' main characteristics. The studies were published between 2000 and 2018, with the vast majority (N = 7) in the 2010s. All were carried out in the United States, except for Averill et al. (2018), in Canada. Settings were diversified, with a prevalence of interventions carried out in the participants' natural environment (N = 4): that is, in these studies, both monitoring and delivery of the incentives were done remotely, without participants transit. Regarding the recruitment process, in half of the studies, patients were invited at the treatment site, while in the other half, recruitment was done through advertisements. The average sample size was 60.3 (SD = 51.5), ranging from 30 to 191 participants. Most participants were men, ranging from 37% to 100% in each study sample. All participants were adults, with a mean age of 40.3 (SD = 6.2) years. Seven studies used standardized instruments to make the diagnosis, with DSM IV being the most common (N = 5). Half of the studies did not describe whether or not comorbidities were present, but those that did reported a high prevalence of associated problems, especially other SUD and mood disorders.

All RCTs employed some biological-marker outcome measure for alcohol use during treatment: six used breathalyzers (two of them with remote monitoring via cell phone), two used transdermal tests, and two included urine tests. In addition, all used self-report as a complementary outcome measure. The biological markers measurement frequency was directly related to the marker used, with the most frequent collection being done every 30 min (with a transdermal test), and the least frequent being done once a week (with a breathalyzer).

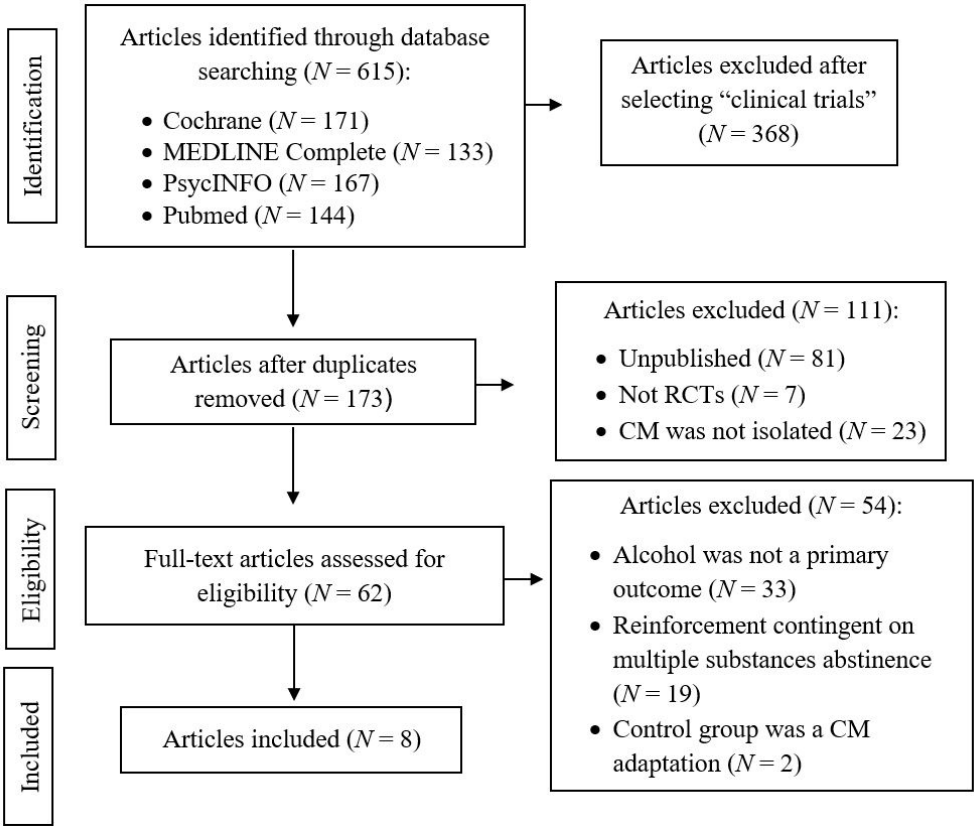

Figure 1. Flow diagram of the literature review process (according to PRISMA)

In relation to control groups, half of the RCTs involved non-contingent reinforcement, followed by reinforcement contingent on the submission of biological marker samples (N = 2). Regarding the CM reinforcement schedule, four RCTs used money, two used vouchers, and the other two, other types of incentives. Five studies used increasing continuous reinforcement as a reinforcement schedule, while the other three used variable-ratio reinforcement, following the fishbowl method developed by Petry (2000). Four RCTs had treatments lasting eight or more weeks; and four lasted for four weeks or less. Half of the studies performed follow-up evaluation.

Table 1
Characterization of the Randomized Clinical Trials Included in this Systematic Review (Ordered by Publication Date)

	Petry et al. (2000)	**Alessi & Petry (2013)**	**Hagedorn et al. (2013)**	**McDonell et al. (2016)**	**Barnett et al. (2017)**	**Averill et al. (2018)**	**Koffarnus et al. (2018)**	**Orr et al. (2018)**
Setting	Outpatient	Natural environment	Outpatient	Outpatient	Natural environment	Natural environment	Natural environment **	University
Recruitment	In site	Advertisement	In site	In site	Advertisement	Invitation letter	Advertisement	Advertisement
Sample	42 men	30 (36.7% men)	191 (97.4% men)	79 (63.5% men)	30 (53.3% men)	37 men	40 (70% men)	34 (63.8% men)
Diagnosis	DSM IV	DSM IV	Medical records	DSM IV	DSM IV	AUDIT	DSM V	DSM IV
Comorbidities	Yes	---	Yes	Yes	---	---	---	Yes
Outcome Measures	Breathalyzer, self-report	Breathalyzer (remote), self-report	Breathalyzer, self-report	Breathalyzer, urine test, self-report	Transdermal testing, self-report	Transdermal testing, self-report	Breathalyzer (remote), self-report	Breathalyzer, urine tests, self-report
Biological Markers Measurement Frequency	Once per day/once per week *	3 times per day	2 times per week	3 times per week	Every 30 min	Every 30 min	3 times per day	Once per day
Control Group	12-step	Reinforcement contingent upon breathalyzer result	Standard treatment	Non-contingent reinforcement	Non-contingent reinforcement + feedback	Reinforcement contingent upon test use + feedback	Non-contingent reinforcement	Non-contingent reinforcement
Reinforcement Schedule	Prize drawing; variable ratio	Vouchers; continuous reinforcement	Vouchers; variable ratio	Gift cards; variable ratio	Money; continuous reinforcement	Money; continuous reinforcement	Money; continuous reinforcement	Money; continuous reinforcement
Duration (weeks)	8	8	8	12	3	4	3	4
Follow-up duration	No follow-up	No follow-up	12 months	3 months	1 month	No follow-up	1 month	No follow-up

Note. --- Data missing in the original article.
*For the first four weeks (intensive care), monitoring took place daily. In the last four weeks, it took place only weekly.
**Participants went to the laboratory only three times (before, after and 1 month after treatment) to participate in evaluation sessions.

CM Efficacy during Treatment

Table 2 presents the RCTs outcomes during treatment and follow-up assessments. Of seven studies that assessed the percentage of abstinence days, five had statistically significant results in favor of CM (Alessi & Petry, 2012; Hagedorn et al., 2013; Koffarnus et al., 2018; McDonell et al., 2016; Orr et al., 2018) and one, a marginally significant favorable result (Barnett et al., 2017). Only the pilot study by Averill et al. (2018) found no significant differences. All five studies that assessed the longest continuous abstinence sequence showed statistically significant differences in favor of CM (Alessi & Petry, 2012; Barnett et al., 2017; Hagedorn et al., 2013; McDonell et al., 2016; Petry et al., 2000). Of the three studies that assessed episodic heavy drinking, two had statistically significant results in favor of CM (McDonell et al., 2016; Petry et al., 2000). Although the pilot study by Barnett et al. (2017) did not find a statistical difference, the effect size was considered medium, suggesting that the lack of significance may be linked to the low sample size. CM was also significantly more efficacious in promoting retention in treatment in the two studies that measured this outcome (Hagedorn et al., 2013; Petry et al., 2000). Finally, three studies (Alessi & Petry, 2021; Barnett et al., 2017; Hagedorn et al., 2013) presented effect sizes for one or more of the analyzed outcomes, with the Cohen coefficient (d) ranging between 0.47 and 0.85 (i.e., moderate to large).

CM Efficacy during Follow-Up

Of the four studies that presented follow-up results for the percentage of abstinence days (Barnett et al., 2017; Hagedorn et al., 2013; Koffarnus et al., 2018; McDonell et al., 2016), three found statistically significant results in favor of CM (Hagedorn et al., 2013.; Koffarnus et al., 2018; McDonell et al., 2016) (see Table 2). It is noteworthy that Hagedorn et al. (2013) found significant differences in favor of CM in this outcome in follow-ups of two, six and 12 months. Only McDonell et al. (2016) evaluated episodic heavy drinking days during follow-up, observing positive and statistically significant differences in favor of CM. All follow-up outcomes were assessed by self-report.

Table 2
Results of Randomized Clinical Trials Included in this Systematic Review (Ordered by Publication Date)

	Petry et al. (2000)	**Alessi & Petry (2012)**	**Hagedorn et al. (2013)**	**McDonell et al. (2016)**	**Barnett et al. (2017)**	**Averill et al. (2018)**	**Koffarnus et al. (2018)**	**Orr et al. (2018)**
During Treatment								
Percentage of abstinence days		$d = 0.62$ $p < 0.01$	$d = 0.54$ $p < 0.001$	$B = 8.29$ $p < 0.05$	$d = 0.74$ $p = 0.053$	Non-significant^	$OR = 9.4$ $x^2 = 26.34$, $p < 0.0001$	$OR = 4.07$ $p < 0.05$
Longest abstinence sequence	$x^2 = 4,5$ $p < 0.05$	$d = 0.52$ $p < 0.01$	$p < 0.001$	$F = 5.55$ $p < 0.05$	$d = 0.85$ $p = 0.031$			
Reduction of heavy alcohol consumption	$x^2 = 5.2$ $p < 0.05$			$B = 6.43$ $p < 0.05$	$d = 0.51$ $p = 0.18$			
Treatment retention	$x^2 = 16.2$ $p < 0.001$		$d = 0.47$ $p < 0.001$					
Last Follow-Up Evaluation								
Percentage of abstinence days			$\beta = 0.35$,[a] $p < 0.05$ $\beta = 0.31$,[b] $p < 0.05$ $\beta = 0.24$,[c] $p < 0.05$	$B = 7.40$ $p < 0.05$	$d = 0.44$ $p = 0.26$		$p = 0.003$	
Reduction of heavy alcohol consumption				$B = 5.69$ $p < 0.05$				

Notes. ^ Result not shown in the original study; [a]After 2 months; [b]After 6 months; [c]After 12 months.

Methodological Quality

Table 3 presents the methodological quality assessment for the studies, based on the general risk of bias criteria proposed by Higgins et al. (2011). Five RCTs had low overall risk, and three had some risk. In the Randomization Process Domain, all had low risk. Orr et al. (2018) did not describe the type of randomization used, but the authors reported (S. McPherson, personal communication, May 1, 2020) that stratified and permuted block randomization was performed, thus presenting a low risk of bias. In the Deviations from Intended Interventions Domain, only Hagedorn et al. (2013) and Petry (2000) presented some risk for not describing intention-to-treat analyzes even with dropouts during treatment. In the Data Lack Domain, all presented low risk. The study by Barnett et al. (2017) excluded missing data from 23.3% of the sample, but due to technical problems in the use of the outcome measure (i.e., the transdermal test bracelet) and not to the result itself. As for the Outcome Measurement Domain, all RCTs used adequate and identical measurements for the control and experimental groups, thus presenting a low risk of bias. In the Reported Outcome Selection Domain, only Averill et al. (2018) presented some risk for not reporting the results of the intergroup comparison of all primary outcomes evaluated.

Table 3
Bias Risk Assessment (General and for each Domain) for the Studies Included in the Review (Ordered by Publication Date)

Domain	Petry et al. (2000)	Alessi & Petry (2012)	Hagedorn et al. (2013)	McDonell et al. (2016)	Barnett et al. (2017)	Averill et al. (2018)	Koffarnus et al. (2018)	Orr et al. (2018)
D1	+	+	+	+	+	+	+	+
D2	-	+	-	+	+	+	+	+
D3	+	+	+	+	+	+	+	+
D4	+	+	+	+	+	+	+	+
D5	+	+	+	+	+	-	+	+
General	-	+	-	+	+	-	+	+

Notes. D1 = Randomization process; D2 = Deviations from intended interventions; D3 = Missing outcome data; D4 = Outcome measurement; D5 = Selection of the reported result. + = Low risk; - = Some risk.

DISCUSSION

Despite solid evidence of CM efficacy and effectiveness for the treatment of several SUD (e.g., Davis et al., 2016; Lussier et al., 2006; Prendesgast et al., 2006), the first systematic CM application in Brazil occurred only a few years ago years (Miguel et al., 2016, 2017, 2018, 2019) and there is no recent systematic review specifically on CM applied to AUD (cf. Gao et al., 2018; Wong et al., 2008). To reduce this knowledge gap, this systematic review analyzed RCTs to assess the efficacy of CM in promoting abstinence, reducing heavy alcohol consumption, and retention in treatment.

The eight RCTs included in the review had a total of 483 participants with AUD, with most samples being composed manly by males—as expected given that men drink more and have more alcohol-related disorders than women (WHO, 2018). The fact that half of them did not describe the presence of comorbidities prevented the analysis of their mediating effects on treatment response. Future studies should control this variable. The analyzed RCTs are generally recent and with samples limited to North America. Although investigations on CM efficacy for SUD started in the 1970s (Higgins et al., 2007; Miller, 1975), the vast majority of RCTs evaluating CM applied specifically to AUD was carried out only in the last decade. In addition, seven of the eight RCTs analyzed were conducted in the United States, a country that has prioritized the adoption of evidence-based practices in treatment services for various diagnoses, including AUD (e.g., NIDA, 1998). Therefore, it is essential to develop evaluation studies of CM for AUD in other countries in order to determine the generality of its effects in different contexts.

The choice of abstinence as the studies primary outcome is justified by the fact that substance use is the only problem behavior in AUD that is essential for the disorder maintenance (i.e., if the user stops using alcohol, it is impossible for them to continue presenting AUD). In addition, abstinence is an outcome that can be more objectively assessed, through biological markers, favoring the reliability of the results obtained (Petry, 2011). The breathalyzer is the oldest and least expensive instrument, and that's probably why it was the most used. Yet, unlike transdermal and urine tests, it has the disadvantage of measuring alcohol use for a restricted period of time (only up to 6 h after alcohol consumption), which can result in inadequate reinforcement being provided

in the face of false positives (Hagedorn et al., 2013). In the study by Koffarnus et al. (2018), for instance, some participants reported having used some alcohol that was not detected by breathalyzer assessments, although these assessments were made three times a day. This limitation is partly overcome by the use of transdermal testing, as it provides several records per day without the participant having to go to a specific location (Barnett et al., 2017). A limitation of the transdermal test, however, is that, in order to detect substance use, it is necessary that transdermal alcohol levels and concentration curves meet specific criteria, which may allow some drink levels to go undetected (Barnett et al., 2017). Future studies should analyze the costs and benefits of each of these measures. We also suggest the standardization of the outcomes included in RCTs, since, in addition to complicating comparisons, this heterogeneity prevents conducting meta-analyses, for example.

The most used CM reinforcement schedule was the increasing continuous reinforcement. That is, participants were reinforced immediately after the target behavior was emitted and the magnitude of this reinforcement gradually increased as the participant remained abstinent. This result connects to the strong association between impulsiveness and substance use (Matta et al., 2014; Mellis et al., 2017), with impulsiveness relating to delay discounting, that is, the current value of a given reward decreases with an increasing interval to receive it (Kirby & Petry, 2004). Delay discounting is especially pronounced in people with SUD (Kirby et al., 1999), who are less sensitive to long-term consequences, hence the importance of immediate reward following target behavior and of the gradual increase in its value as the individual remains abstinent.

The prevalence of non-contingent reinforcement as a procedure used in control groups is justified by the importance of verifying whether the dependent variable change was caused by the incentive presentation itself or by the incentive being contingent on the target behavior (or, more strictly, to products related to the target behavior, verified through objective measures)—a critical aspect of the CM. The researchers' preference for using money and vouchers as reinforcers is due to its function as generalized reinforcers, applicable to any study setting or type of participant (Petry, 2011).

All RCTs that measured abstinence during treatment, with a single exception, found statistically significant results favorable to CM (vs. the control group). Similar results were found for days of episodic heavy drinking and treatment retention—although these outcomes have been evaluated in few studies. When effect sizes were reported, these ranged between medium and large effect sizes (Cohen, 1988), pointing not only to their statistical but also clinical relevance. These results clearly demonstrate the positive effects of CM for the treatment of AUD and corroborate the positive results of previous reviews that had investigated the application of this intervention to SUD in general (e.g., Davis et al., 2016; Lussier et al., 2006; Prendesgast et al., 2006). The only RCT that did not show statistically significant results for the primary outcomes was a pilot study, with the smallest sample among the analyzed RCTs (30 participants divided into two groups) and that evaluated the use of transdermal testing as an outcome measure (Averill et al., 2018). Averill et al. (2018) raised the hypothesis that the use of this measure alone would be sufficient to obtain the desired results. That is, the constant use of the bracelet on which the test is installed would be the main factor responsible for the change in the direction of abstinence. Unfortunately, neither of the two studies that used the transdermal test used a control group without a bracelet as a comparison, which prevents us from confirming this hypothesis.

As for the effects after the treatment, again only one study (out of four) could not find statistically significant effects favorable to CM. This study, by Barnett et al. (2017), as the research by Averill et al. (2018), was also a pilot study, with the smallest sample among the RCTs and evaluating the transdermal test. This demonstrates CM efficacy also during follow-up (Gao et al., 2018). These assessments, however, were performed in only half of the RCTs analyzed, suggesting the importance of continuing the investigation of the CM effects after the end of treatment.

The analyzed RCTs had very good methodological quality, with five studies presenting low risk of general bias and none presenting high risk. The very nature of the studied intervention prevents the professionals involved in the studies and participants from being blinded to the current experimental condition. However, the participants in the control groups also went through the recruitment process and were exposed to some form of treatment or intervention aimed at alcohol use, just like the participants in the experimental groups. Furthermore, the RCTs reported no deviations from CM implementation. These two elements led to eight RCTs being categorized as having low risk of bias in the blinding component of the Deviations from Intended Interventions Domain. In addition, all RCTs included used an objective measure to assess the outcome (i.e., breathalyzer, transdermal or urine test), an uncommon procedure in research on psychosocial interventions, in which the exclusive use of self-report instruments prevails (McLeod, 2003). Although CM positive results for AUD have been documented since the first publication in the area (Miller, 1975), this seminal study had serious methodological problems, such as the use of different outcome measures between control and experimental groups. Errors like this, which would result in a high risk of bias, did not occur in any of the studies included in this review, which attests to the methodological development of CM studies over the years.

As for its limitations, this review excluded unpublished studies (e.g., theses), studies using a methodological design other than RCT, studies that evaluated reinforcement contingent on multi-substance abstinence, and studies in which CM was integrated with other interventions. The

inclusion of these other studies could help answer new questions, such as: is the reinforcement of alcohol abstinence alone more effective in reducing its use than the reinforcement of abstinence from multiple substances? And, in the case of choosing a single substance to reinforce abstinence, what criteria should be considered? Orr et al. (2018), for example, studying alcohol and tobacco users, found that reinforcing tobacco abstinence more markedly decreased the use of both substances than reinforcing alcohol abstinence. In a future systematic review of the literature, an expansion of the inclusion criteria for studies may allow us to answer questions like these.

CONCLUSION

AUD represents a serious public health problem in Brazil, requiring efforts to disseminate and implement scientifically-based practices for this disorder. This review contributes to the advancement of evidence-based practices in Psychology, being the first systematic review to assess CM efficacy specifically for AUD. The results provide robust evidence of CM efficacy for the treatment of AUD. We suggest that this practice be widely disseminated, and that empirical studies (RCTs, single-subject experiments and case studies) involving CM applied to AUD be conducted in Brazil, especially in public services such as Psychosocial Care Centers - Alcohol and Drugs [*Centros de Atenção Psicossocial Álcool e Drogas*], in order to assess the efficacy of this intervention in the Brazilian population.

REFERENCES

Abdalla, R. R., Massaro, L., Miguel, A. D. Q. C., Laranjeira, R., Caetano, R., & Madruga, C. S. (2018). Association between drug use and urban violence: Data from the II Brazilian National Alcohol and Drugs Survey (BNADS). *Addictive Behaviors Reports*, *7,* 8-13. https://doi.org/10.1016/j.abrep.2017.11.003

*Alessi, S. M., & Petry, N. M. (2013). A randomized study of cellphone technology to reinforce alcohol abstinence in the natural environment. *Addiction, 108*(5), 900-909. https://doi.org/10.1111/add.12093

American Psychiatric Association. (2014). *Manual diagnóstico e estatístico de transtornos mentais: DSM 5*. Artmed.

Andreuccetti, G., De Carvalho, H. B., de Carvalho Ponce, J., De Carvalho, D. G., Kahn, T., Muñoz, D. R., & Leyton, V. (2009). Alcohol consumption in homicide victims in the city of São Paulo. *Addiction*, *104* (12), 1998-2006. https://doi.org/10.1111/j.1360-0443.2009.02716.x

*Averill, F., Brown, T. G., Robertson, R. D., Tchomgang, A., Berbiche, D., Nadeau, L., & Ouimet, M. C. (2018). Transdermal alcohol monitoring combined with contingency management for driving while impaired offenders: A pilot randomized controlled study. *Traffic Injury Prevention, 19*(5), 455-461. https://doi.org/10.1080/15389588.2018.1448079

*Barnett, N. P., Celio, M. A., Tidey, J. W., Murphy, J. G., Colby, S. M., & Swift, R. M. (2017). A preliminary randomized controlled trial of contingency management for alcohol use reduction using a transdermal alcohol sensor. *Addiction, 112*(6), 1025-1035. https://doi.org/10.1111/add.13767

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2a ed.). Lawrence Erlbaum Associates.

Davis, D. R., Kurti, A. N., Skelly, J. M., Redner, R., White, T. J., & Higgins, S. T. (2016). A review of the literature on contingency management in the treatment of substance use disorders, 2009–2014. *Preventive Medicine, 92*, 36-46. https://doi.org/10.1016/j.ypmed.2016.08.008

de Carvalho Ponce, J., Muñoz, D. R., Andreuccetti, G., de Carvalho, D. G., & Leyton, V. (2011). Alcohol-related traffic accidents with fatal outcomes in the city of Sao Paulo. *Accident Analysis & Prevention*, *43*(3), 782-787. https://doi.org/10.1016/j.aap.2010.10.025

Degenhardt, L., Charlson, F., Ferrari, A., Santomauro, D., Erskine, H., Mantilla-Herrara, A., Whiteford, H., Leung, J., Naghavi, M., Griswold, M., Rehm, J., Hall, W., Sartorius, B., Scott, J., Vollset, S. E., Knudsen, A. K., Haro, J. M., Patton, G., Kopec, J., ... Vos, T. (2018). The global burden of disease attributable to alcohol and drug use in 195 countries and territories, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Psychiatry*, *5*(12), 987-1012. https://doi.org/10.1016/S2215-0366(18)30337-7

Diehl, A., Cordeiro, D. C., & Laranjeira, R. (2019). *Dependência química: Prevenção, tratamento e políticas públicas* (2a ed.). Artmed.

Gao J., Cao J., Guo T., & Xiao Y. (2018). Association between alcoholic interventions and abstinence rates for alcohol use disorder: A meta-analysis. *Medicine, 97*(50). https://doi.org/10.1097/MD.0000000000013566

*Hagedorn, H. J., Noorbaloochi, S., Simon, A. B., Bangerter, A., Stitzer, M. L., Stetler, C. B., & Kivlahan, D. (2013). Rewarding early abstinence in Veterans Health Administration addiction clinics. *Journal of Substance Abuse Treatment, 45*(1), 109-117. https://doi.org/10.1016/j.jsat.2013.01.006

Higgins, J. P. T., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., Savović, J., Schulz, K. F., Weeks, L., & Sterne, J. A. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*, *343*, d5928. https://doi.org/10.1136/bmj.d5928

Higgins, S. T., Silverman, K., & Heil, S. H. (2007). *Contingency management in substance abuse treatment*. Guilford Press.

Kendall, J. (2003). Designing a research project: Randomised controlled trials and their principles. *Emergency Medicine Journal, 20*(2), 164. https://doi.org/10.1136/emj.20.2.164

Kirby, K., & Petry, N. (2004). Heroin and cocaine abusers have higher discount rates for delayed rewards than alcoholics or non-drug-using controls. *Addiction*, *99*(4), 461-471. https://doi.org/10.1111/j.1360-0443.2003.00669.x

Kirby, K., Petry, N., & Bickel, W. (1999). Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls. *Journal of Experimental Psychology: General*, *128*(1), 78-87. https://doi.org/10.1037//0096-3445.128.1.78

*Koffarnus, M. N., Bickel, W. K., & Kablinger, A. S. (2018). Remote alcohol monitoring to facilitate incentive-based treatment for alcohol use disorder: A randomized trial. *Alcoholism: Clinical and Experimental Research, 42*(12), 2423-2431. https://doi.org/10.1111/acer.13891

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*(2), 192-203. https://doi.org/10.1111/j.1360-0443.2006.01311.x

Massaro, L. T. D. S., Adesse, L., Laranjeira, R., Caetano, R., & Madruga, C. S. (2019). Estupros no Brasil e relações com o consumo de álcool: Estimativas baseadas em autorrelato sigiloso. *Cadernos de Saúde Pública*, *35*, e00022118. https://doi.org/10.1590/0102-311x00022118

Matta, A., Gonçalves, L. F., & Bizarro, L. (2014). Desvalorização pelo atraso, dependência química e impulsividade. *Avances en Psicología Latinoamericana, 32*(2), 217-230. https://doi.org/10.12804/apl32.2.2014.03

Mellis A. M., Woodford A. E., Stein J. S., & Bickel W. K. (2017). A second type of magnitude effect: Reinforcer magnitude differentiates delay discounting between substance users and controls. *Journal of the Experimental Analysis of Behavior*, *107*, 151-160. https://doi.org/10.1002/jeab.235

McCrady, B. S. (2016). Transtornos por uso de álcool. In D. H. Barlow (Org.), *Manual clínico dos transtornos psicológicos: Tratamento passo a passo* (5a ed., pp. 531-583). Artmed.

*McDonell, M. G., Leickly, E., McPherson, S., Skalisky, J., Srebnik, D., Angelo, F., Vilardaga, R., Nepom, J. R., Roll, J. M., & Ries, R. K. (2017). A randomized controlled trial of ethyl glucuronide-based contingency management for outpatients with co-occurring alcohol use disorders and serious mental illness. *American Journal of Psychiatry, 174*(4), 370-377. https://doi.org/10.1176/appi.ajp.2016.16050627

McLeod, J. (2003). Tests, rating scales and survey questionnaires. In J. McLeod (Ed.), *Doing counseling research* (2a ed., pp. 55-70). Sage Publications.

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. (2016). Contingency management is effective in promoting abstinence and retention in treatment among crack cocaine users in Brazil: A randomized controlled trial. *Psychology of Addictive Behaviors, 30*(5), 536-543. https://doi.org/10.1037/adb0000192

Miguel, A. Q. C., Madruga, C. S., Cogo-Moreira, H., Yamauchi, R., Simões, V., Ribeiro, A., da Silva, C. J., Fruci, A., McDonell, M., McPherson, S., Roll, J. M., & Laranjeira, R. R. (2017). Contingency management targeting abstinence is effective in reducing depressive and anxiety symptoms among crack cocaine-dependent individuals. *Experimental and Clinical Psychopharmacology, 25*(6), 466–472. https://doi.org/10.1037/pha0000147

Miguel, A. Q. C., Madruga, C. S., Simões, V., Yamauchi, R., Silva, C. J. D. A., Abdalla, R. R., McDonell, M., McPherson, S., Roll, J. M., Mari, J. J., & Laranjeira, R. R. (2018). Crack cocaine users views regarding treatment with contingency management in Brazil. *Substance Abuse Treatment Prevention, and Policy*, *13*(1), 1-6. https://doi.org/10.1186/s13011-018-0144-7

Miguel, A. Q. C., Madruga, C. S., Simões, V., Yamauchi, R., Silva, C. J. D., McDonell, M., McPherson, S., Roll, J., Laranjeira, R. R. & Mari, J. D. J. (2019). 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. *Psicologia: Reflexão e Crítica*, *32*. https://doi.org/10.1186/s41155-019-0127-2

Miller, P. M. (1975). A behavioral intervention program for chronic public drunkenness offenders. *Archives of General Psychiatry, 32*(7), 915-918. https://doi.org/10.1001/archpsyc.1975.01760250107012

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine*, *151*(4), 264-269. https://doi.org/10.7326/0003-4819-151-4-200908180-00135

National Institute on Drug Abuse. (1998). *A community reinforcement plus values approach: Treating cocaine addiction.* (Therapy Manuals for Drug Addiction).

*Orr, M. F., Smith, C. L., Finlay, M., Martin, S. C., Brooks, O., Oluwoye, O. A., Leickly, E., McDonell, M., Burduli, E., Barbosa-Leiker, C., Layton, M., Roll, J. M., & McPherson, S. M. (2018). Pilot investigation: Randomized-controlled analog trial for alcohol and tobacco smoking co-addiction using contingency management. *Behavioural Pharmacology, 29* (5), 462-468. https://doi.org/10.1097/FBP.0000000000000379

Petry, N. M. (2011). *Contingency management for substance abuse treatment: A guide to implementing this evidenced-based practice*. Routledge.

*Petry, N., Martin, B., Cooney, J. L., & Kranzler, H. R. (2000). Give them prizes, and they will come: Contingency management for treatment of alcohol dependence. *Journal of Consulting and Clinical Psychology, 68*(2), 250-257. https://doi.org/10.1037/0022-006X.68.2.250

Pilling, S., Strang, J., Gerada, C., & NICE (2007). Psychosocial interventions and opioid detoxification for drug misuse: Summary of NICE guidance. *BMJ*, *335*(7612), 203–205. https://doi.org/10.1136/bmj.39265.639641.AD

Ponce, J. D. C., Andreuccetti, G., Jesus, M. D. G. D. S., Leyton, V., & Muñoz, D. R. (2008). Álcool em vítimas de suicídio em São Paulo. *Archives of Clinical Psychiatry* (São Paulo), *35*, 13-16. https://doi.org/10.1590/S0101-60832008000700004

Prendergast, M., Podus, D., Finney, J., Greenwell, L., & Roll, J. (2006). Contingency management for treatment of substance use disorders: A meta-analysis. *Addiction, 101*(11), 1546-1560. https://doi.org/10.1111/j.1360-0443.2006.01581.x

Wong C. J, Silverman K., & Bigelow G. E. (2008). Alcohol. In S. T. Higgins, K. Silverman, & S. H. Heil (Eds.), *Contingency management in substance abuse treatment* (pp. 120-139). Guilford Press.

World Health Organization. (2018). *Global Status Report on Alcohol and Health 2018*. https://www.who.int/publications-detail/global-status-report-on-alcohol-and-health-2018

The studies included in this systematic review are highlighted with an asterisk (*)