CampbellCollaboration

Social Welfare Coordinating Group

12-step programs for reducing illicit drug use

Martin Bøg, Trine Filges, Lars Brännström, Anne-Marie Klint Jørgensen and Maja Karrman Fredrikksson

A Campbell Systematic Review
2017:2

Published: February 2017
Search executed: September 2016

CampbellCollaboration

The Campbell Library comprises:

- Systematic reviews (titles, protocols and reviews)
- Policies and Guidelines Series
- Methods Series

Go to the library to download these resources, at:

www.campbellcollaboration.org/library/

Better evidence for a better world

Colophon

Title	12-step programs for reducing illicit drug use
Institution	The Campbell Collaboration
Authors	Bøg, Martin Filges, Trine Brännström, Lars Jørgensen, Anne-Marie Klint Fredriksson, Maja Kärrman
DOI	10.4073/csr.2017.2
No. of pages	149
Last updated	February 2017
Citation	Bøg M, Filges T, Brännström L, Jørgensen AMK, Fredriksson MK. 12-step programs for reducing illicit drug use: a systematic review. Campbell Systematic Reviews 2017:2 DOI: 10.4073/csr.2017.2
ISSN	1891-1803

Roles and responsibilities	Bøg, Filges, Brännström, and Jørgensen contributed to the writing and revising of this review. The search strategy was developed by Jørgensen and Frederiksson. Martin Bøg will be responsible for updating this review as additional evidence accumulates and as funding becomes available.
Editors for this review	Editor: Cathy Bennett, Brandy Maynard Managing editor: Jane Dennis, Catriona Shatford
Sources of support	SFI Campbell / Danish Social Research Institute
Declarations of interest	The authors have no vested interest in the outcomes of this review, nor any incentive to represent findings in a biased manner.

Corresponding author	Martin Bøg SFI Campbell/ Danish National Centre for Social Research Herluf Trollesgade 11 DK-1052 Copenhagen K Denmark E-mail: martin.bog@gmail.com

Full list of author information is available at the end of the article.

Campbell Systematic Reviews

The Campbell Collaboration was founded on the principle that systematic reviews on the effects of interventions will inform and help improve policy and services. Campbell offers editorial and methodological support to review authors throughout the process of producing a systematic review. A number of Campbell's editors, librarians, methodologists and external peer reviewers contribute.

The Campbell Collaboration
P.O. Box 4404 Nydalen
0403 Oslo, Norway
www.campbellcollaboration.org

Table of contents

Plain language summary

12-step programs for reducing illicit drug use are neither better nor worse than other interventions

Illicit drug abuse has serious and far-reaching implications for the abuser, their family members, friends, and society as a whole. Preferred intervention programs are those that effectively reduce illicit drug use and its negative consequences, and are cost-effective as well. Current evidence shows that overall, 12-step programs are just as effective as alternative, psychosocial interventions. The costs of programs are, therefore, an important consideration. However, the strength of the studies is weak and further evidence regarding the effectiveness of 12-step programs is needed.

What did the review study?

Illicit drug abuse is a globally recognised problem leading to high human, social and economic costs.

What is the aim of this review?
This Campbell systematic review examines the effectiveness of 12-step programs in reducing the use of illicit drugs. The review summarises findings from 10 studies, nine of which were conducted in the United States.

The 12-step program, modelled on the approach of Alcoholics Anonymous and adopted by Narcotics Anonymous and others, aims for complete abstinence. The 12-step approach is used both by self-help groups and for professional treatment called Twelve Step Facilitation (TSF).

This review examines the effectiveness of 12-step programs in reducing the use of illicit drugs. Secondary outcomes considered are on criminal behaviour, prostitution, psychiatric symptoms, social functioning, employment status, homelessness, and treatment retention.

What studies are included?

Included studies assess 12-step interventions for participants with illicit drug dependence using randomized controlled trials and quasi-experimental studies. Study populations are participants who have used one or more types of illicit drugs, regardless of gender and ethnic background.

A total of 10 studies consisting of 1,071 participants are included in the final evaluation. Nine of the studies were conducted in the United States, and one in the United Kingdom. The studies compare the 12-step program to alternative interventions. Nine studies were included in meta-analysis.

What are the main results in this review?

There is no difference in the effectiveness of 12-step interventions compared to alternative psychosocial interventions in reducing drug use during treatment, post treatment, and at 6- and 12-month follow-ups. 12-step programs combined with additional treatment did have a significant effect at 6-month follow-up, but this finding is based on few studies and is not found at 12-month follow-up.

There is some evidence that 12-step programs retain fewer of their participants than other programs, but the evidence has shortcomings. No effect was found on other secondary outcomes.

What do the findings in this review mean?

The main evidence presented in this review suggests that 12-step programs for reducing illicit drug use are neither better nor worse than other interventions.

This conclusion should be read with caution given the weakness of the evidence from the studies. The power to detect a difference between the 12-step interventions and alternative psychosocial interventions was low and the estimated effect sizes were small. Many studies failed to adjust for the fact that the intervention is administered to groups, and so may overestimate effects. Given all these shortcomings, further evidence regarding the effectiveness of this type of intervention, especially in self-help groups, is needed.

How up to date is this review?

The review authors searched for studies published until September 2016. This Campbell Systematic Review was published in February 2017.

What is the Campbell Collaboration?

The Campbell Collaboration is an international, voluntary, non-profit research network that publishes systematic reviews. We summarise and evaluate the quality of evidence for social and economic policy, programs and practice. Our aim is to help people make better choices and better policy decisions.

Executive summary

BACKGROUND

The effects of substance dependence have serious implications for the individual, the family and friends of the substance dependent individual, and society at large. Practitioners and public health policy makers have an interest in finding effective treatments that are also cost-effective. This review examined the effectiveness of 12-step programs aimed at illicit drug dependent participants compared to no intervention, treatment as usual, and other interventions.

OBJECTIVES

The main objective of this review was to systematically evaluate and synthesise effects of 12-step interventions for participants with illicit drug dependence against no intervention, treatment as usual, and alternative interventions. The primary outcome of interest was drug use. Secondary outcomes of interest comprised criminal behaviour, prostitution, psychiatric symptoms, social functioning, employment status, homelessness and treatment retention.

SEARCH METHODS

An extensive search strategy was used to identify studies meeting inclusion criteria. We searched electronic bibliographic databases in January 2010, October 2011, July 2013, August 2015, and September 2016. Searches for this review were performed on multiple international and Nordic databases. In total 11 databases were searched including PsycInfo, SocIndex, and Medline. A substantial range of grey literature sources were searched including governmental repositories, targeted web sites and trial registers. We checked the reference lists of primary studies, hand-searched relevant key journals, and searched the Internet using Google and Google Scholar. We also contacted researchers who had published in the area of 12-step interventions. Neither language nor date restrictions were applied to the searches. The conclusions of this review are based on the most recent searches performed September 2016.

SELECTION CRITERIA

Studies had to meet the following criteria in order to qualify for inclusion in the review:

- Intervention - only studies that considered 12-step interventions were eligible for inclusion.
- Study Design - only studies using a RCT/QRCT design or a QES with a well-defined control group were eligible for inclusion.

- Comparison - studies that compared 12-step to either no intervention or to other interventions were eligible for inclusion.
- Participants - only studies where the drug of choice of participants was an illicit drug (established either by self-report or via clinician) were eligible for inclusion. Where only a subset of study participants were illicit drug users, a study was only eligible if it reported outcomes separately for the subgroup of illicit drug users.

DATA COLLECTION AND ANALYSIS

Descriptive and numerical characteristics of included studies were coded by one review author. A second review author independently checked coding, and any disagreements were resolved by consensus. We used an extended version of the Cochrane Risk of Bias tool to assess risk of bias of included studies. One review author evaluated the risk of bias of all included studies. A second review author independently checked the assessment and disagreements were resolved by consensus. Random-effects meta-analysis was used to synthesise effect sizes. We compared 12-step to other interventions, and 12-step with add-on to other interventions with the same add-on. For each comparison we conducted separate meta-analyses by time: during treatment, at treatment end, and at 6- and 12-month follow-up. Sensitivity of the results to risk of bias was assessed. Publication bias was assessed by the use of funnel plots.

MAIN RESULTS

The total number of potentially relevant records was 21,974 (database search: 17,416, grey literature search: 2,639, hand search and others: 1,919), of these 428 records were screened in full text. Thirteen reports met the inclusion criteria, with six reports contributing data on three independent studies. In total 10 studies were included in the review.

Seven of the included studies used a RCT design, two studies used a QRCT design, and one study used a QES design. One study, assessed as high risk of bias, was excluded from data synthesis. Thus, nine studies with a total of 1,071 participants contributed data to the analyses. These nine studies all considered outpatient settings where interventions were manual-based and delivered by trained therapists. In seven studies, treatment was partially or fully delivered in group therapy sessions. The reported statistical analyses were not corrected for this design element.

Seven studies contributed data to the comparison of 12-step intervention to alternative psychosocial interventions during treatment, at treatment end, and at 6-and 12-month follow-up. The seven studies did not all contribute data to all time points. Analyses did not reveal any statistically significant differences, for the primary outcome of drug use, between 12-step and the alternative set of interventions.

Three studies contributed data to the comparison of 12-step intervention with an add-on to alternative psychosocial interventions with an add-on. Drug use was assessed during treatment, post treatment, and at 6- and 12-months follow-up. All studies did not contribute data to all time points. We found no statistically significant effect size estimates during and post treatment. We found statistically significant effect size estimates at 6-month follow-up favouring 12-step with an

add-on compared to alternative interventions with add-on (Hedges' g =0.48, 95% CI: 0.06 to 0.90, and g=0.45, 95% CI: 0.03 to 0.88). No statistically significant effect size estimates were found at 12-months follow-up.

There was no strong indication of heterogeneity between studies (I^2 did not exceed 75%). Results were robust to sensitivity analysis, and there was no observed evidence of publication bias.

AUTHORS' CONCLUSIONS

The results of this review suggest that 12-step interventions to support illicit drug users are as effective as alternative psychosocial interventions in reducing drug use.

This conclusion should be seen against the weight of evidence. A total of seven studies contributed data to analyses comparing 12-step interventions and alternative psychosocial interventions. The power to detect differences was low, and estimated effect sizes were small. In addition most studies delivered treatment as group therapy, but did not correct the analysis for the dependence between participants assigned to the same group.

Only one study reported results of the effects of self-help group attendance on drug use. This study was excluded from synthesis following the risk of bias assessment. Given the preponderance with which self-help 12-step interventions are delivered in practice, further evidence regarding the effectiveness of this type of intervention is needed.

1 Background

1.1 DESCRIPTION OF THE CONDITION

Illicit drug production and use[1] remains a severe problem worldwide (United Nations Office on Drugs and Crime, UNODC, 2010). A central issue in reducing the worldwide drug problem is the demand for illicit drugs, and hence the need to identify effective methods for reducing their use.

Prescription and recreational drugs should be differentiated from one another. In this review, we reserve the term 'drug use' to apply to the illegal, nonmedical use of drugs. Globally, the United Nations Office on Drugs and Crime (UNODC) estimates that between 155 and 250 million people (3.5 to 5.7 percent of the population aged 15-64) used illicit substances at least once in 2008 (UNODC, 2010). Illicit substances include opium/heroin (opiates), coca/cocaine, cannabis[2], and amphetamine-type stimulants (including MDMA/Ecstasy). Cannabis is the most commonly used illicit substance (with an estimated 129 – 190 million users worldwide), followed by amphetamine-group substances, cocaine and opiates. UNODC considers some types of drug use to be more problematic than others, and defines problem drug use as that which involves the injection of drugs or the long-duration/regular use of opioids, cocaine and/or amphetamines. For 2008, UNODC estimates that between 16 and 38 million people worldwide are problematic drug users (UNODC, 2010).

Drug use is linked to a range of health and social problems, including crime, prostitution, and homelessness (Office of National Drug Control Policy, ONDCP, 2000; Shelton, Taylor, Bonner, & van den Bree, 2009; Silbert, Pines, & Lynch, 1982). The European Monitoring Centre for Drugs and Drug Addiction estimates that drug-induced deaths account for approximate 4 percent of all deaths of Europeans aged 15-39 (European Monitoring Centre for Drugs and Drug Addiction, EMCDDA, 2010). A number of studies have attempted to estimate the social costs of drug use, in terms of both the direct cost and the indirect costs of drug use (EMCDDA, 2010). For Finland these costs have been estimated at between EUR 200 million and EUR 300 million in 2007, and in Scotland at EUR 5.1 billion in 2006.

The high human, social and economic costs of illicit drug use motivate the strong political interest in treatment for illicit drug use and in identifying effective treatments. The main types of treatment are cognitive-behavioural therapies, motivational enhancement, contingency management,

[1] The terms use, misuse, abuse and dependence will be used interchangeably throughout the review and refer to an addiction stage of drug usage.

[2] Cannabis is illegal in most, but not all countries. For example, use of cannabis in small amounts is not a criminal offence in the Netherlands.

psychoanalysis, network therapy and – the object of this review –12-step programs (Galanter & Kleber, 2008).

1.2 DESCRIPTION OF THE INTERVENTION

12-step approaches to the treatment of drug use are widespread in many countries. The oldest and most widely attended 12-step groups are provided by Alcoholics Anonymous (AA) that began in 1935 and have more than 2 million members according to their own recent membership survey (Alcoholics Anonymous, 2012; Mäkelä et al., 1996). The principles of AA serve as a model for other 12-step programs, of which Narcotics Anonymous (NA) is the largest focusing on drug use (Narcotics Anonymous, 2012). Today NA has more than 58,000 weekly meetings in 131 countries (Narcotics Anonymous, 2010). NA accepts all individuals with drug addiction, regardless of the particular drug or combination of drugs used. Other 12-step groups restrict themselves to specific types of abuse, such as Cocaine Anonymous, Pills Anonymous and Marijuana Anonymous (Cocaine Anonymous, 2012; Pills Anonymous, 2012; Marijuana Anonymous, 2012). The stated objective of the 12-step approaches is complete abstinence from the use of drugs, whereas other treatments such as psychosocial interventions or opioid substitution may focus on reducing drug use (EMCDDA, 2010). 12-step treatment approaches assume that, as a result of biological and/or psychological vulnerability, individuals have lost control over their drug use. Treatment attempts to bring about the individual's acceptance of the disease model of addiction, (i.e. that addiction is a lifelong disease), of an "addict" identity, and of abstinence as a treatment goal. It also attempts to motivate involvement in 12-step activities (for example, attending meetings, obtaining a sponsor, working through the 12 steps) (Finney, Noyes, Coutts, & Moos, 1998). The core ideology of the approach is to offer individuals a new way of living that will support them in breaking the cycle of addiction and in maintaining abstinence (Mercer & Woody, 1999). The suggested prescription for abstinence, referred to as the "six pack", is: *don't use no matter what, go to meetings, ask for help, get a sponsor, join a group, and get active* (Laudet, 2008).

The 12-step self-help groups work to specific principles. The meetings typically adhere to a prescribed format including 12-step readings (The Preamble, How and Why, The 12 Steps) at the start of the meeting, and a reciting of the Serenity prayer at the end for individuals who wish to do so (Laudet, 2008). The disease model of addiction is central to the 12-step philosophy and recovery, and is seen as being a significant part of the process of attaining and maintaining abstinence. Recovery is viewed as a lifelong process, and members thus regard themselves as "recovering" (Mercer & Woody, 1999).

The basic idea is that individuals work their way into recovery by going through "12 Steps" starting with the recognition of being addicted to drugs, also known as "hitting the bottom" [3], and ending with the capability of helping others out of their own addiction (Narcotics Anonymous, 2008). A central element in the 12-step self-help groups is that participation is voluntary and that recovering individuals, and not professional staff, guide the treatment. Another important component is

[3] The insistence on hitting bottom lies in the belief that few individuals will be sincerely motivated to commit to recovery unless they have "hit the bottom". This bottom can be wherever the individual allows it to be. The central idea is that individuals must come to a turning point where they accept that they have reached a stage of complete defeat to drugs (Alcoholics Anonymous, 2005).

sponsorship whereby a member who has made progress in the 12-step recovery program shares his or her experience on an individual and continuous basis with another member who is attempting to attain or maintain abstinence through the 12 steps (Straussner & Spiegel, 1996; Laudet, 2008). Sponsors share their own "experience, strength and hope" with the sponsees and accompany them in working the steps towards recovery. The idea is that sponsorship also helps oneself in maintaining abstinence, formulated as: "*the cardinal virtue of sponsorship is the momentary loss of self-centeredness*" (Jennings & Alcoholics Anonymous, 1990). Sponsorship is something that a member her/himself decides to become involved in. In addition, sponsors themselves have their own sponsors who help them in their own struggle for abstinence (Narcotics Anonymous, 2008).

The exact wording of the 12 steps differs slightly between groups – the 12 steps of NA are presented in section 10.1 (Narcotics Anonymous, 2008). The steps contain a strong spiritual emphasis and encourage members to look outside themselves for strength (to seek a higher power) and to embrace spiritual values and practices that are outlined in the 12 steps. "*A power greater than ourselves*" is mentioned in step 2 and "*God as we understand Him*" in step 3. In addition, the steps emphasise the importance of reconstructing relationships with people who have been harmed by the drug use (e.g., family members). Inherent in the 12 steps is the realisation that addiction is a disease and as such is beyond personal influence. However an individual can decide to change and oppress disease. The strong spiritual emphasis is unique to the 12 steps and the texts that are used include wording that appear religious. On the other hand, the texts do not endorse a particular faith and the "*God as we understand Him*" phrase is open to interpretation. The spiritual emphasis may mean that some drug users are opposed to the intervention for ideological reasons, and hence the claim made by some that 12-step only works for individuals with the right motivation and outlook on life (Fiorentine, 1999).

The "12 Steps" are accompanied by "12 Traditions" providing guidelines for the self-help groups (Narcotics Anonymous, 2008). The 12 traditions of Narcotics Anonymous are presented in section 10.1. Because the individual groups are autonomous, there can be differences between them, but the basic concepts are the same. Each group meets at a regular time and place, and is in principle open to all drug users – the only requirement being the wish to become "clean". Attendance is decided individually, but beginners are encouraged to attend "90 meetings in 90 days". Furthermore, participation is not time-limited and the time needed to do the steps is also decided individually (Straussner & Spiegel, 1996). Often, drug users who have succeeded in attaining abstinence continue to participate in the meetings for years, and continue working the steps day after day. Abstinence anniversaries or birthdays are considered major accomplishments and an important way to mark success in attaining abstinence.

In addition to the self-help groups, 12-step approaches are used in Twelve Step Facilitation (TSF) treatments (MATCH, 1997; Ries, Galanter, & Tonigan, 2008). These are typically of limited duration and organised around a treatment facilitation centre staffed by professionals, with treatment based on 12-step principles. In short, TSF is the integration of 12-step self-help groups with professional treatment. Usually, the individual will participate in 12-step meetings (NA or other) after completing TSF treatment, and a parameter of success for the TSF treatment is that the individual becomes motivated and ready for the self-help group. The best known TSF treatment is the "Minnesota model", originating from three centres founded in Minnesota in the late 1950s

(Cook, 1988). The Minnesota model is characterised by the use of the 12-step philosophy as a foundation for therapeutic change, where the treatment goal is total abstinence. TSF treatment can be delivered in both inpatient and outpatient settings; the duration of the treatment can vary, but is typically around 12-24 weeks.

A cardinal rule of both TSF and self-help groups is anonymity. In attempt to protect individuals from society's stigmatisation, the 12-step approach gives priority to preserving members' anonymity. Anonymity inevitably poses a challenge to research.

In this review, the focus is on 12-step treatments for users of illicit drugs, and we will include both treatments based on self-help groups working with the 12 steps (like NA) and TSF treatments. Since drug users may also be alcohol dependent (Kessler et al., 1997), we will include studies where alcohol misuse is present provided drug abuse is the key drug of choice of participants. Thus, although alcohol can be part of the substance abuse we do not consider 12-step treatments primarily dealing with and aimed at treating alcohol dependency.

1.3 HOW THE INTERVENTION MIGHT WORK

The 12-step interventions included in this review are aimed at supporting the substance users to refrain completely from or reducing their use of illicit drugs. Abstinence is achieved by the drug user through working his/her way through the 12 steps. A central issue that dates back to the AA tradition is the acknowledgement of the addiction, and the acceptance of support by a sponsor in a self-help group, or by professionals in a TSF setting. One of the keys to success posed by the AA, and hence the NA, is "*the therapeutic value of addicts working with other addicts*" along with the cardinal idea that the 12 steps offer "*a design for living*", a way of learning to live, that teaches skills and helps individuals to navigate and reach recovery (Alcoholics Anonymous, 2005; Narcotics Anonymous, 2008). The steps are carefully organised in an order, starting with the basic skills and continuing to the more advanced changes, that individuals should gradually seek to integrate into their lives. Having a sponsor and being a sponsor is an important part of 12-step self-help groups. Sponsorship is viewed as an important tool in the process of recovery in that it helps to grasp the components of living, offers encouragement and support such as when relapse occurs, but it may also "kill" any complacency among the sponsors themselves, and thereby help them sustain their self-monitoring (Hornbacher, 2010). Also, central to the NA program is the statement that spirituality mediates 12-step involvement and later abstinence. It is suggested that, by working the 12 steps, one will have a spiritual awakening, and that continued practice of spiritual principles will lead to sustained abstinence (Narcotics Anonymous, 2008).

The benefits associated with involvement in 12-step programs, and the mechanisms by which these benefits occur, have been thoroughly explored over the past two decades. Relatively little is known, however, about which specific behaviours catalyse the therapeutic psychological mechanisms. Self-efficacy, or the confidence to remain abstinent, has been identified as a major component and a consistent predictor of subsequent improvement (Moos & Timko, 2008; MATCH, 1997). The importance of spirituality for later abstinence is currently unclear (Maude-Griffin et al., 1998; Moos & Timko, 2008; Tonigan & Connors, 2008). Regarding sponsorships, Humphreys & Noke (1997) point out that this social network component of treatment can be more effective in helping

the substance user than the support from concerned family members. The reliance on positive reinforcement (e.g., by recognising abstinence anniversaries) and behavioural modelling (e.g., by having a sponsor) have also been proposed as an underlying mechanism of change (Morgenstern, Bux, & Labouvie, 2002; Witkiewitz & Marlatt, 2011). Membership demographics for Narcotics Anonymous collected at the 2009 World Convention of NA in Barcelona, Spain reveal that more men than women are members (58% vs. 42%); only very few teenagers are members of NA (2%), whereas the most typical member is aged 41-50 years (34%); ethnicity is dominated by Caucasians (73%) and most members are employed (71%) (Narcotics Anonymous, 2010). Perhaps due to this profile, Fiorentine (1999) notes that the 12 steps have been argued to work best for Christian, white, middle-class males. According to Fiorentine (1999) studies fail to support this view. Fiorentine (1999) also accentuates that 12-step interventions may also be inappropriate for drug users with major psychiatric disorders, drug users in early stages of addiction, and drug users uncomfortable with the religious or spiritual emphasis.

Treatment setting may also affect models of treatment and participant characteristics. In inpatient settings, patients stay at the treatment facility overnight and possibly for extended periods of time. Treatment typically includes a first period of detoxification followed by initial intensive treatment, including preparing patients for returning to community-based care settings. Patients participating in treatments in outpatient settings typically stay in their own home, while attending treatment at a treatment facility. Due to the nature of treatment in- and out-patient settings may differ substantially in participant characteristics such as e.g. substance abuse severity, clinical symptoms, consequences from use, motivation, and labour market attachment.

The criminal justice system in the US is responsible for a substantial proportion of referrals to community-based treatment programs. Legal referral may consists of a probation officer's recommendation to enter treatment, the choice in a drug court between jail time and treatment, the requirement of a judge to enter treatment as a precondition for probation, etc. (Farabee, Prendergast, & Anglin, 1998). The degree of legal pressure that the patient faces to comply with treatment may affect his/her motivation for change. Intrinsic motivation, understood as the patient's willingness to change his/her substance use, has been linked as key to treatment success (e.g. Prochaska, & DiClemente, 1982). As such, court-mandated treatment may be expected to lead to less desirable treatment outcomes since enrolment is due to legal pressure and not intrinsic motivation (Farabee, Prendergast, & Anglin, 1998). Even patients who enter "voluntarily" may face pressure in the form of social pressure from e.g. family and friends (Perron, & Bright, 2008). Perhaps surprisingly, mandated treatment patients show substance abuse outcomes similar to and sometimes better than voluntary treatment patients (Kelly, Finney, & Moos, 2005). This may partly be explained by patient characteristics. Kline (1997) found that legally referred patients were younger, involved fewer African-Americans, less substance use, and fewer drug-related health problems compared to voluntary patients in residential treatment. Kelly, Finney, & Moos (2005) examined treatment retention and treatment outcomes at Veteran Affairs residential treatment facilities, where a main program component was 12-step group involvement. The justice system involved (JSI) patients had lower severity of substance abuse, fewer consequences of abuse, were younger, more likely to be white, and were less motivated for treatment than those who were not justice system involved in the same program. JSI patients, particularly JSI patients who were

mandated to treatment, showed similar or better outcomes for abstinence and remission at 1 and 5 years post treatment, even after controlling for pre-treatment differences.

1.4 WHY IT IS IMPORTANT TO DO THE REVIEW

Illicit drug use has significant costs for individuals and societies in terms of social, health and criminal problems. 12-step programs are one of the most widespread treatments for drug use internationally. The general belief among clinicians is that 12-step is an effective approach (Forman, Bovasso, & Woody, 2001). In the US, for example, it is common procedure by courts to mandate 12-step treatment. Although a large number of studies have examined the use of the programs, no systematic knowledge of the effectiveness of the intervention is currently available. This is, in part, due to the strict anonymity policy and the insistence on fluid membership, especially in the self-help groups, which makes it difficult for researchers to track members.

A Cochrane review evaluated the effectiveness of 12-step programs on alcohol dependence. Ferri, Amato, & Davoli (2006) provided a narrative synthesis of the effectiveness of 12-step approaches (like AA or TSF) compared with alternative psychosocial interventions (e.g. cognitive behavioural therapy, motivational enhancement therapy, relapse prevention). They concluded that 12-step approaches were as effective as alternative approaches in terms of improving drinking consequences, and reduction of alcohol consumption. Ferri and colleagues did not explicitly exclude co-dependence on illicit drugs, nor did they report whether participants, in the eight trials that were included in the review, had comorbid substance dependence. Treatment effectiveness for comorbid participants treated for alcohol disorder might inform the effectiveness for participants included in the present review. Even so, prevalence of alcohol disorder only in the US was 7.35 percent, the prevalence of any drug use disorder was 2.00 percent, and the comorbidity of alcohol and (any) drug use disorder was 1.10 percent (2001-02 National Epidemiologic Survey on Alcohol and Related Conditions; Stinson, Grant, Dawson, Ruan, Huang, & Saha, 2005). This suggests that while comorbidity is not uncommon in the population, there is substantial non-overlap between populations. As such, one should be cautious in extrapolating results from Ferri, Amato, & Davoli (2006) to the population of interest to this review. Furthermore, individuals dependent on drugs, particularly illicit drugs, are often forced into a lifestyle that differs considerably from individuals who are dependent on alcohol, due to the criminal aspects of drug use. People who are drug dependent are often engaged in illegal activity when obtaining their substance of dependence and/or the capital needed for its procurement. Research also suggests that it may be more difficult for the drug user to benefit from the 12 steps compared to the individual who is dependent on alcohol, possibly due to the particular impact of drugs on brain neurons (Laudet, 2008).

With its broad applicability, minimal cost, and potential benefit, the 12-step approach has great appeal to policy makers. Knowledge about the effectiveness of the method compared to other treatments, as well as knowledge about the effect of different program elements, will therefore be of considerable interest to policy makers and practitioners.

2 Objectives

The objective of this review was to assess the effectiveness of 12-step programs to reduce illicit drug use. The following questions were addressed to determine the effectiveness of 12-step programs aimed at illicit drug users:

1. reducing illicit *drug use*, during treatment, at treatment end, and at follow-up, compared with no intervention or a competing psychosocial intervention, and when used in conjunction with either a pharmaceutical add-on or another psychosocial add-on?

2. reducing *criminal behaviour and prostitution*, during treatment, at treatment end, and at follow-up, compared with no intervention or a competing psychosocial intervention, and when used in conjunction with either a pharmaceutical add-on or another psychosocial add-on?

3. reducing *psychiatric symptoms*, during treatment, at treatment end, and at follow-up, compared with no intervention or a competing psychosocial intervention, and when used in conjunction with either a pharmaceutical add-on or another psychosocial add-on?

4. improving *social functioning*, during treatment, at treatment end, and at follow-up, compared with no intervention or a competing psychosocial intervention, and when used in conjunction with either a pharmaceutical add-on or another psychosocial add-on?

5. improving *employment status*, during treatment, at treatment end, and at follow-up, compared with no intervention or a competing psychosocial intervention, and when used in conjunction with either a pharmaceutical add-on or another psychosocial add-on?

6. reducing *homelessness*, during treatment, at treatment end, and at follow-up, compared with no intervention or a competing psychosocial intervention, and when used in conjunction with either a pharmaceutical add-on or another psychosocial add-on?

7. improving *treatment retention*, during treatment, at treatment end, and at follow-up, compared with no intervention or a competing psychosocial intervention, and when used in conjunction with either a pharmaceutical add-on or another psychosocial add-on?

3 Methods

3.1 TITLE REGISTRATION AND REVIEW PROTOCOL

The title for this systematic review was approved by The Campbell Collaboration on 16 October 2010. The review protocol was approved on 2 September 2013. The title registration and protocol are available at: https://www.campbellcollaboration.org/library/12-step-programmes-illicit-drug-abuse-reduction.html

3.2 CRITERIA FOR INCLUDING STUDIES IN THE REVIEW

The purpose of this review was to synthesise the best available evidence on the effects of 12-step programs, broadly understood as participation in self-help programs or manual based programs designed to reduce illicit drug use and delivered in in- or out-patient settings. No restrictions on the language or the publication status of studies were applied. When we could not determine whether a report met the inclusion criteria we attempted to contact investigators. We applied a time limit of 14 days from sending our inquiry. If the study investigators notified us before this deadline that they would be able to provide the information in a short time, we included the information. If investigators did not reply to our inquiry in time, the study was listed as "Awaiting classification". All inquiries and answers were stored electronically.

The following criteria were used to select studies eligible for synthesis.

3.2.1 Types of study designs

Study designs eligible for inclusion were:

- Randomised Controlled Trials (RCTs) – studies where individual participants, or group of participants were randomised to control or treatment conditions, including trials that randomly assigned individual participants to group treatment.
- Quasi-randomised controlled trials (QRCTs) - where participants were allocated by, for example, alternation, birth date, date of the week, case number or alphabetical order, to treatment or control conditions.
- Quasi-experimental studies (QESs) – studies where participants were assigned to control and treatment conditions in a non-random manner with a control group where pre-treatment group equivalence is demonstrated via matching, statistical controls, or where there is evidence of equivalence on key risk variables (see section 3.4.2), or where key risk variables are controlled for statistically.

We justified the inclusion of QRCTs and QESs because the open-door membership policy of 12-step programs generally, and for the self-help groups in particular, may pose considerable problems in assessing the effectiveness of treatment using a RCT design. In addition, studies that utilise quasi-experimental designs may produce as efficient and unbiased estimates of intervention effects as studies utilising a RCT design (Shadish & Cook, 2009).

3.2.2 Types of participants

The population included participants who used illicit drugs, regardless of age, gender or ethnic background. We included participants who used one or more types of illicit drugs.

Many studies included participants with both alcohol and drug use. Studies where alcohol use dominated drug use, either in consumption or in severity as measured by, for example, the Addiction Severity Index (McLellan, Luborsky, Woody, & O'Brien, 1980), were excluded. Studies that included participants who had both participants with illicit drugs or alcohol as their drug of choice were only included if they reported outcomes separately for participants with illicit drugs as their drug of choice. When the reported information on usage was insufficient for a judgement to be made, we contacted the study authors for clarification and used this information to determine whether the study should be included.

We included studies of individuals who were enrolled in 12-step treatment regardless of the way in which their problem was labelled, and we regarded the terms 'use', 'abuse', 'misuse' or 'dependence' of illicit drugs as equivalent.

3.2.3 Types of settings

The review included studies in which treatment was delivered in inpatient, outpatient, or self-help groups.

3.2.4 Types of interventions

The review included studies evaluating 12-step interventions (in either the self-help or TSF format) delivered with the explicit aim of stopping or reducing illicit drug use. The following core principles were present in the studies with 12-step intervention:

- Addiction was viewed as an illness.
- There was a theme of spirituality (for example, a belief in a higher power).
- The individual discussed problems within a fellowship of peers trying to help and encourage one another.
- General guidance was provided in the "12 Steps".

We included studies evaluating interventions that used the 12-step principles regardless of setting (for example, inpatient or outpatient) or the duration of treatment. 12-step interventions that focused solely on treating alcohol dependency, such as AA programs, were excluded even though the study participants may also have been addicted to illicit drugs.

3.2.5 Types of comparisons

Eligible comparison conditions were no intervention, a waitlist control condition, or any other intervention(s).

3.2.6 Types of outcomes

Given that 12-step treatment accepts abstinence as the only successful treatment outcome, the primary outcome for this review was abstinence or reduction of drug use as measured by:

- Biochemical tests
- Estimates of drug use

Secondary outcomes were as follows:

- Criminal behaviour
- Prostitution
- Psychiatric symptoms
- Social functioning
- Employment status
- Homelessness
- Retention

Primary outcome measures might be reported in the form of urine toxicology screens, as self-reports of drug use (or reported by others such as parents, caregivers, or therapists) either as a prevalence measure, or a measure of use in the past month such as the Timeline Follow Back instrument (Sobell et al., 1996), or the relevant portion of the Addiction Severity Index (McLellan et al., 1980). Studies were only included if they considered one of the primary outcomes.

Secondary outcomes, such as criminal behaviour and prostitution, may be self-reported, such as from the Addiction Severity Index sub-components (criminal and legal), registers or files. The Symptom Checklist-90-R (SCL-90-R) (Derogatis, 1983) or a similar validated scale may be used to report outcomes relating to psychiatric outcomes. Social functioning may be measured by the Social Functioning Questionnaire (SFQ) (Tyrer et al., 2005) or a similar validated scale. Employment and homelessness might come from self-report, or registers. Finally, retention could be measured in a variety of ways, such as the study authors' own conception of treatment completion, number of sessions attended, or percentage of sessions attended. Types of Time Points

Outcomes were considered at the following intervals:

- During treatment. All studies that provided numerical effect sizes for synthesis were psycho-social manual-based interventions delivered by professional or trained therapists. Many studies provided primary outcome measures while participants were being treated.
- Post treatment/treatment end. Post treatment measures were taken directly after the end of intervention, or shortly thereafter.
- Follow-up. Follow-up outcome measures were taken between 1 and 18 months after the end of treatment.

This classification of relevant time periods follows the convention frequently reported in the literature and differs slightly from that outlined in the published protocol (see section 8.2 for further details).

3.3 SEARCH METHODS FOR IDENTIFICATION OF STUDIES

3.3.1 Electronic searches

Relevant studies were identified through electronic searches of bibliographic databases, government policy databanks and internet search engines. No date or language restrictions were applied to the searches. All databases where searched in the original search in January 2010. Access to some databases has changed throughout the four updated searches to this review. The date of the last search was September 19, 2016. The following bibliographic databases were searched:

- ASSIA - (Searched through ProQuest) - Searched until July 2013 – no access for further years
- Cochrane Library (including CENTRAL) - http://onlinelibrary.wiley.com/cochranelibrary/search?searchRow.searchOptions.searchProducts=clinicalTrialsDoi - Searched until 19. September 2016
- Embase (Searched through OVID host) - Searched until 19. September 2016
- Medline (Searched through OVID host) - Searched until 19. September 2016
- PsycINFO (Searched through EBSCO host) - Searched until 19. September 2016
- CINAHL (Searched through EBSCO host) - Searched until July 2013 - no access for further years
- Science Citation Index (Searched through ISI Web of Science) - Searched until 19. September 2016
- Social Science Citation Index (Searched through ISI Web of Science) - Searched until 19. September 2016
- SocINDEX (Searched through EBSCO host) - Searched until September 2016
- Sociological Abstracts (Searched through ProQuest) - Searched until 19. September 2016
- Dissertation Abstracts (Searched through ProQuest) – Searched until 19. September 2016

The following Nordic library databases were searched:

- BIBSYS - http://www.bibsys.no/en/ - In the newest update of the review, the access to BIBSY (now Oria) was limited. The documented search strategy in section 9.1 reflects searches until August 2015
- Bibliotek.dk - https://bibliotek.dk - Searched until 16. September 2016
- DiVA (Digitale vetenskapeliga arkivet) - http://www.diva-portal.org/smash/search.jsf?dswid=7660 - Searched until July 2013.
- LIBRIS - http://libris.kb.se/ - Searched until 16. September 2016
- SweMed+ - http://svemedplus.kib.ki.se/Default.aspx?searchform=advanced Searched until 16. September 2016.
- Artikelsök - http://artikelsok.se/ Searched until July 2013– no access for further years.

3.3.2 Search terms

Exact search strategies for each database can be found in section 9.1.

3.3.3 Searching other resources

We examined the reference lists from relevant reviews identified in the electronic searches, and from included primary studies for studies that potentially met inclusion criteria. In addition, international experts were contacted to attempt to identify unpublished and on-going studies.

3.3.4 Grey literature

The search strategy for the grey literature search was based on the search string for the electronic database search. Due to the limited search capacity on grey literature information resources, web pages and search engines, a shortened search string was used. An example of the search strategies used to identify grey literature and google searches can be found in section 9.1. The most recent search for grey literature was performed September 6, 2016. Following websites and resources were searched for relevant grey literature:

- Canadian Evaluation Society - http://evaluationcanada.ca/ - Searched 3. September 2016.
- NARCIS (National Academic Research and Collaborations Information System) - http://www.narcis.nl/about/Language/en – Searched 3. September 2016.
- Government of Canada - https://www.canada.ca/en/index.html - Searched 3. September 2016.
- USA.gov - https://www.usa.gov/ - Searched 3. September 2016.
- Australian Government - http://www.australia.gov.au/about-government/publications - Searched 3. September 2016.
- Ministry of Social Affairs and the Interior - http://sim.dk/publikationer.aspx - Searched 3. September 2016.
- Government Offices of Sweden - http://www.government.se/ - Searched 4. September 2016.
- Government.no - https://www.regjeringen.no/en/id4/ - Searched 4. September 2016.
- European Union - https://europa.eu/european-union/index_en - Searched 4. September 2016.
- Theses Canada Portal - http://www.bac-lac.gc.ca/eng/services/theses/Pages/theses-canada.aspx - Searched 4. September 2016.
- National Library of Germany - http://www.dnb.de/EN/Home/home_node.html - Searched 4. September 2016.
- Social Care Online - http://www.scie-socialcareonline.org.uk/ - Searched 4. September 2016.
- DART-Europe E-theses Portal - http://www.dart-europe.eu/basic-search.php - Searched 4. September 2016.

- Information for Practice - http://ifp.nyu.edu/archive/ - Searched 4. September 2016.
- Open Grey - http://www.opengrey.eu/ - Searched 5. September 2016.
- National Institute on Drug Abuse - https://www.drugabuse.gov/ - Searched 5. September 2016.
- European Monitoring Centre for Drugs and Drug Addiction - http://www.emcdda.europa.eu/ - Searched 5. September 2016.
- Sbustance Abuse and Mental Health Services Administration - http://www.samhsa.gov/data/node/20 - Searched 5. September 2016.
- NCJRS (National Criminal Justice Reference Service) - https://www.ncjrs.gov/index.html - Searched 6. September 2016.

Additional searches were conducted using Google and Google Scholar, and the first 200 hits were examined in each case.

3.3.5 Hand searching

Searching was performed on journal editions from January 2010 to September 2013 and finally from August 2015 to September 2016 in attempt to identify any published studies that may not have been found in the electronic search. The most recent hand search was performed September 26, 2016. The following five international journals was hand searched for relevant studies:

- *Addiction* – Searched 26. September 2016.
- *Journal of Consulting and Clinical Psychology* - Searched 26. September 2016.
- *Journal of Substance Abuse Treatment* - Searched 26. September 2016.
- *Journal of Clinical and Adolescent Psychology* - Searched 26. September 2016.
- *Research on Social Work Practice* - Searched 26. September 2016.

3.3.6 Selection of studies

The screening process was executed in two separate phases.

Phase 1: Two review team members independently screened each title and abstract obtained from the search procedures for inclusion. Each reviewer coded each citation according to the pre-specified inclusion criteria. This information was stored in a Reference Manager database. Disagreements were handled by discussion and consensus agreement.

Points of discussion included whether the design of the study met the inclusion criteria, and whether outcomes reported were consistent with the focus of the present review. The decisions available to the reviewer were: (1) 'In' (include for full article scan), (2) 'Unclear' (include for full scan), and (3) 'Out' (citation eliminated).

Citations that met the initial inclusion criteria were retrieved for full review using available library resources.

Phase 2: Two review team members independently screened the full articles for inclusion. As with the previous procedure, the studies were screened against the inclusion criteria, with results tracked in an Excel database. If the citation was excluded at this stage, the reviewer provided a brief description of the reason for dismissal. When there was a disagreement, two reviewers discussed the citation and reached an agreement.

3.4 DATA COLLECTION AND ANALYSIS

3.4.1 Data extraction and management

Study level data such as author, year, and report type were extracted. In addition intervention characteristics, such as duration, intensity, type of delivery (e.g. group or individual), fidelity, outcome assessors, were coded. Most of this information was used in the risk of bias assessment of

each study. Outcome measurements were also extracted. If relevant effect sizes could be extracted directly, this information was coded. If the report did not provide effect sizes directly, reviewers extracted other information, such as the *t*-statistic and the sample size, which would allow effect sizes to be calculated. If sufficient information was not provided in the report, this was coded, and provided the basis for the detailed data requests made to the report authors. The full codebook is available in section 7.2 of the protocol.

One reviewer extracted descriptive and numerical data from the included studies. A second reviewer checked all coded information for accuracy, and in case of discrepancies, reviewers would jointly agree on the final coding. The data were coded and stored electronically in Microsoft Excel.

When descriptive or numerical data were not available in the published reports or only partly available we contacted the authors, requesting the required information. Eight reports comprising seven corresponding authors contained insufficient information for relevant effect sizes to be extracted. In line with the protocol for the review, we contacted the authors, requesting the information and allowed a two-week deadline[4] from the initial attempt to contact the corresponding author. One author team, comprising two reports, responded positively to our enquiries and provided the necessary missing data. Thus for these reports data were complete; indeed, we were able to code outcomes that were not available in the published reports.

3.4.2 Assessment of risk of bias in included studies

We assessed the methodological quality of studies using a risk of bias model developed by Prof. Barnaby Reeves in association with the Cochrane Non-Randomised Studies Methods Group[4]. This model is an extension of the Cochrane Collaboration's risk of bias tool; it covers risk of bias both for RCTs, and risk of bias for non-randomised studies with a well-defined control group. The point of departure for the risk of bias model is the Cochrane Handbook for Systematic Reviews of Interventions (Higgins & Green, 2008). The risk of bias model for non-randomised studies is an elaboration of the existing Cochrane risk of bias tool and incorporates particular attention to selection bias and risk of confounding. The extended tool includes assessment of risk of bias on a 5-point scale for some items.

3.4.2.1 Risk of bias judgement items

The risk of bias model is based on nine items (see 9.4.1). For some items, risk is assessed to be High, Low, or Uncertain; other items are judged on a 5-point scale where 1 corresponds to No/Low risk of bias and 5 correspond to Yes/High risk of bias. A score of 5 indicates that the risk of bias is sufficiently high that the findings will not be considered in the data synthesis (because they are more likely to mislead than inform).

The nine risk of bias items concern **sequence generation** (relevant for selection bias), **allocation concealment** (relevant for selection bias), **confounders** (relevant for selection bias in non-randomised studies), **blinding** (relevant for performance, detection and attrition bias),

[4] This risk of bias model was introduced by Prof. Reeves at a workshop on risk of bias in non-randomized studies at SFI Campbell, February 2011. The model is a further development of work carried out in the Cochrane Non-Randomised Studies Method Group (NRSMG). See also Reeves, Deeks, Higgins, & Wells (2011).

incomplete outcome data (relevant for attrition bias), **selective outcome reporting** (relevant for reporting bias), **other potential threats to validity** (relevant for performance, detection and other sources of bias), **a priori protocol,** and **a priori analysis plan** (relevant for reporting bias).

3.4.2.2 Confounders

An important part of the risk of bias assessment for a non-randomised study is how the confounding factors have been dealt with (see 9.4.2.2). Selection bias is understood as systematic baseline differences between groups, which can compromise comparability. Baseline differences can be observable (e.g., age and gender) and unobservable to the researcher (e.g. motivation and "ability"). There is no single non-randomised study design that resolves the selection problem in all circumstances. Different designs attempt to solve the problem under different assumptions and require different types of data, particularly in relation to factors that are unobservable. The "right" method depends on the assumptions about the nature of the process by which participants are selected into a program. As there is no universally correct way to construct counterfactuals, we assessed the extent to which the identifying assumptions (the assumption that makes it possible to identify the counterfactual) were explained and discussed by the study investigators.

In this review, the risk of bias from confounding is an additional item for each non-randomised study, and were assessed for each outcome. Such an assessment requires a list of pre-specified potential confounders. We identified the following confounding factors as the most relevant:

- Age
- Gender
- Socio-economic status
- Mental health problems
- History of drug use

The motivation for focusing on these confounders was that they are major risk factors related to drug use. Young people have a higher risk of use than older people (Labouvie, 1996), women have lower risk than men and have different drug use patterns (Brady & Back, 2008), and people with poor socio-economic status have higher risk (Spooner & Hetherington, 2004). The issue of drug users with mental health problems needs special attention, because the mental health problems can interfere with the effect of the drug treatment (Ross, 2008). Finally, the history of drug use is important for the likelihood of treatment success, e.g. duration of use and previous treatment (Greenfield & Hennessy, 2008). We also assessed how each study dealt with factors that are unobservable.

The risk of bias item takes into account the following:

- Proportion of confounders considered.
- Whether most important confounders were considered.
- Precision with which confounders were measured.
- Extent of imbalance between groups at baseline.
- Care with which adjustment was done.

The final judgement of this risk of bias item was made on a scale from 1 to 5 (or unclear), where a score of 1 reflects low risk and a score of 5 reflects a high risk of bias in relation to confounding. For a judgement of low risk of bias in this item, all important confounders should be balanced at baseline or measured "well" and "carefully" controlled for in the analysis. The final judgement of the confounding item is included in the overall risk of bias table.

One review author evaluated the risk of bias of all included studies. A second review author independently checked the assessment and disagreements were resolved by consensus.

3.4.3 Effect size calculations

Effect sizes were extracted from each included study by the methods described below. Two effect size measures were used in the review. For continuous measures we calculated effect sizes as standardised mean differences (Hedges' *g*). For dichotomous outcome data we calculated the odds ratio. Where appropriate effect sizes were recoded such that a positive effect size reflected an outcome that favoured the 12-step intervention (see also 4.4).

The primary outcome for the review was drug use. Seven studies (Carroll et al. 1998; Carroll et al., 2012; Higgins et al., 1991; McKay et al., 1997; Petry et al., 2010; Schottenfeld et at., 2011; Wells et al., 1994) reported drug use on a continuous scale. Six studies reported drug use on a discrete scale (Bisset, 2002; Caroll et al., 1998; Higgins et al., 1991; Maude-Griffin et al., 1998; Petry et al., 2010; Schottenfeld et al., 2011). Continuous measures were outcomes such as "days of cocaine use per week", whereas the dichotomous outcomes measured constructs such as "(complete) abstinence" (e.g. Bisset 2002, Maude-Griffin et al., 1998). Because continuous outcome measures were reported in a majority of studies, and because the dichotomous outcome could be considered a dichotomised version of an underlying continuous construct we used the standardised mean difference (Hedges' *g*). By transforming the relevant effect sizes we were able to include them in the same meta-analysis, thereby increasing power to detect possible differences. We transformed effect sizes from log odds ratio to standardised mean difference, for those studies that only reported drug use on a discrete scale. If a study reported both a continuous and a discrete scale measure for the same time point, we used the outcome reported on a continuous scale. Effect sizes were transformed for Bisset (2002) (post and follow-up measure), Maude-Griffin et al. (1998) (during, post and follow-up), Petry et al. (2010) (follow-up), and Schottenfeld et al. (2011) (follow-up).

All secondary outcomes, apart from retention, were reported on a continuous scale. Therefore effect sizes for these outcomes were calculated as standardised mean difference (Hedges' *g*). Five studies (Bisset et al., 2002; Carroll et al., 1998; Carroll et al., 2012; Higgins et al., 1991; Schottenfeld et al., 2011) reported retention as a discrete measure ("treatment completion") and four studies reported retention on a continuous scale, such as "number of sessions attended" (Caroll et al., 2012; McKay et al., 1997; Petry et al., 2010; Wells et al., 2010). Because dichotomous outcome measures were in the majority and we stated "treatment completion" as the outcome measure of interest in the protocol (section 3.1.4) a log odds ratio effect size was chosen. We transformed effect sizes to log odds ratio, for those studies where only a continuous scale effect size for retention was available (McKay et al., 1997; Petry et al. 2010; Wells et al., 1994).

Where appropriate outcomes were recoded such that a more positive score reflected an improvement of the outcome in question. For example Caroll et al. (2012) used self-reports in the Timeline Followback to measure drug use as "number of days per month using cocaine". In this case, using raw outcomes, a superior outcome of 12-step over the alternative intervention would imply a negative effect size. Accordingly these estimated effect sizes were multiplied by -1.

The table in section 11.2 presents outcome measures by study and time of measurement.

3.4.3.1 Discrete data

For discrete measures, the effect size was calculated as the log odds ratio (LOR). The LOR and its approximate standard deviation were calculated as (Lipsey & Wilson, 2001:53-54):

$$LOR = log\left(\frac{ad}{bc}\right), SE_{LOR} = \sqrt{\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}}$$

where *a* is the frequency of "good" outcomes in the treatment group (e.g. the number retained), *b* is the frequency of "bad" outcomes in the treatment group (the number of participants who were not retained), and *c* and *d* are the number of good and bad outcomes in the control group, respectively.

When appropriate (for studies that only reported a discrete measure for drug use at a given time point) we transformed the LOR to Hedges' g using the Cox-transformation (see section 3.4.3.4).

3.4.3.2 Continuous data

For continuous measures, we calculated the effect size as the standardised mean difference (Hedges' g), and applied the small *N* correction. Hedges' (adjusted) g and its standard error are calculated as (Lipsey & Wilson, 2001:47-49)

$$g = \left(1-\frac{3}{4N-9}\right)\times\left(\frac{\bar{X}_1-\bar{X}_2}{s_p}\right), SE_g = \sqrt{\frac{N}{n_1 n_2}+\frac{g^2}{2N}}$$

where $N = n_1 + n_2$ is the total sample size, $\bar{X}$ denotes the (adjusted) mean of a group, and s_p is the pooled standard deviation defined as

$$s_p = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{(n_1-1)+(n_2-1)}}$$

Here, s_1 and s_2 denotes the standard deviation of the two groups.

When data were not available we extracted the effect size from auxiliary statistics. For example Wells, Peterson, Gainey, Hawkins, & Catalano (1994) reported means and the *t*-statistic for retention, but no standard deviation. By using standard techniques (Lipsey & Wilson, 2001) we were able to construct an effect size.

When appropriate (for studies that only reported a continuous measure for retention) we transformed Hedges' g to LOR using the Cox-transformation (see section 3.4.3.4).

3.4.3.3 Data from graphs and figures

When data in the reports were insufficient to construct an effect size, we attempted to contact study authors directly. We only successfully established contact with one group of authors who had retained primary data or summary statistics. Three studies (Higgins et al., 1991; Schottenfeld et al., 2011; Maude-Griffin et al., 1998) provided graphs or figures that permitted the construction of relevant effect sizes, either directly or with additional assumptions. In these instances, and in order to base the meta-analysis on as much data material as possible, we attempted to recover effect sizes from graphs and figures provided. For each of these three studies, two members of the review team (MB and TF) independently took measurements, results were compared, and a consensus measurement agreed upon.

From Figure 1 (p. 1222, top panel) in Higgins, Delaney, Budney, & Bickel, (1991) we constructed a measure of drug use during the intervention based on number of weeks abstinent. Since we knew the total number of participants assigned to each condition, we were able to reconstruct individual durations of abstinence. We used a ruler to convert the percentages reported in the figure to number of participants. From this data, means and standard deviations by condition were constructed and formed the basis for calculating effect sizes.

Schottenfeld, Moore, & Pantalon (2011) did not provide any numerical data that were directly amenable to meta-analysis. Instead, data were constructed from Figures 3 and 4 (drug use during and at 3 months follow-up, respectively). We were able to read precise individual participant measurements from Figure 3. From these measurements, means and standard deviations by condition were constructed. In order to construct measurements of the 3 months follow-up, we used the information provided in Figure 1 about how many participants were followed up, and we used a ruler in order to get an accurate measure of the point prevalence of percent abstinent by condition. We then constructed a 2 x 2 frequency table.

We constructed effect sizes for Maude-Griffin et al. (1998) post treatment and at 14-week follow-up in a similar fashion. Maude-Griffin et al. (1998) did not provide the sample size at each measurement point by condition, but did report overall follow-up rates. Under the assumption that follow-up rates were independent of assignment, we were able get a measure of the "percent of subjects abstinent from cocaine" by taking measurements with a ruler. We constructed the 2 x 2 frequency table from these measurements.

3.4.3.4 Effect size transformations

We used the Cox-transformation to transform continuous scale effect sizes (Hedges' g) to log odds ratio and vice versa. The Cox-transformation for the effect size and the associated standard error is (Sánchez-Meca, Marín-Martínez, & Chacón-Moscoso, 2003):

$$LOR = 1.65 \times g, SE_{LOR} = 1.65 \times SE_g$$

We applied the transformation from LOR to *g* for drug use, and from *g* to LOR for retention, where appropriate.

3.4.3.5 Dependent effect sizes

Dependencies between effect sizes may occur for a multitude of reasons, including when studies are multi-arm trials. For example, Carroll, Nich, Ball, McCance, & Rounsavile (1998) conducted a 5 arm trial. In 3 of these, disulfiram was given as an add-on to the psychosocial interventions: TSF, CBT, and CM. Since the comparisons relevant to this review involve TSF against an alternative, the comparisons between TSF/CBT and TSF/CM are dependent.

A study may also report several effect sizes for the same theoretical construct. For example, Petry, Weinstock, Alessi, Lewis, & Dieckhaus (2010) measured "drug use" during treatment both as longest consecutive number of weeks of negative samples submitted and also as the proportion of negative samples submitted.

Effect sizes may also be serially correlated, such as when a study contributes effect sizes at several time points. Where this occurred, only one outcome measure per construct was retained for meta-analysis.

Data from multi-arm trials were synthesised in different meta-analyses. We also split analyses by time points to avoid dependencies between effect sizes.

3.4.3.6 Unit of analysis issues

If designs other than individually randomised, parallel-group randomised trials were included, we described any methods used to address clustering, matching or other design features.

12-step interventions may be delivered in groups or individually, and often take place with some sessions delivered individually, and some delivered as group sessions. For example, Maude-Griffin et al. (1998) individually randomised participants to receive either CBT or 12-step facilitation delivered as three group therapy sessions and one individual counselling session each week for 12 weeks.

A study design where participants are individually randomised to treatment, but that treatment is delivered in a group setting, are known as *individually randomised group treatment* (IRGT) trials (Pals et al., 2008). The analysis in such a study design must correct for the fact that dependencies may arise between individuals that happen to receive the intervention in the same group. The analogy is the cluster randomised trial where clusters of participants are randomised to treatment. The analysis of cluster randomised trials must correct standard errors for the dependencies among individual participants in clusters. The correction involves knowledge of the intra-cluster correlation coefficient (ICC) and the (mean) group size. With this in hand, the estimated standard errors can be corrected with a *design effect*. Unfortunately none of studies report the ICC, and only 2 studies report group size. In the two studies that reported group size, Wells et al., 1994 delivered to groups with a mean size of 12, while Petry et al. (2010) reported a mean group size of 4 (for TSF intervention) and 4.5 (for the comparison condition). Pals et al. (2008) reviewed 34 IRGT trials in public health and noted that reporting of ICCs were very rare. The data that are available produced

ICC estimates ranging between 0.04 and 0.44 depending on participants, interventions, and outcomes.

Since we had no relevant information on either ICC or group size, we decided that we could not reliably correct estimates for this unit of analysis error. Instead, studies that employed an IRGT design and did not correct the analysis for this design choice were scored on the "Other Bias" item. We then carried out a sensitivity analysis on this item.

3.5 DATA SYNTHESIS

Studies that scored 5 on a risk of bias item were excluded from the meta-analysis.

3.5.1 Effect size synthesis

Separate meta-analyses were carried out by outcome, and by time point (see section 3.2.7). In addition, analyses were organised based upon whether the comparison was 12-step versus other psychosocial intervention, or 12-step + add-on vs other psychosocial intervention + add-on. Only one included study (Bisset, 2002) contributed data to the comparison of 12-step vs no intervention. Accordingly meta-analysis could not be performed. We report the study level effect size for this comparison.

We retained only one effect size measure per comparison for each outcome construct for meta-analysis. When a study contributed dependent effect sizes, these were synthesised in separate meta-analyses. For example, as mentioned in section 3.4.3.5, the study by Carroll et al. (1998) contributed two relevant effect sizes: TSF versus CBT and TSF versus CM; these were synthesised in separate meta-analyses.

All analyses were carried out using inverse variance weighted random effects statistical models that incorporated both the sampling variance and between-study variance components into study level weights. We decided to use a random effects model to represent the overall effect as we expected included studies to deal with diverse populations of participants and intervention types.

Results of meta-analyses were presented by outcome and time of measurement. For each analysis we reported the number of studies, the average effect size with 95% confidence intervals, and where appropriate measures of heterogeneity, and in the case of I^2 with an uncertainty interval (see section 3.5.2).

3.5.2 Assessment of heterogeneity

Heterogeneity was assessed through the use of the χ^2-test. A *p*-value smaller than 0.1 was taken as indication of significant heterogeneity of treatment effects (protocol, sec. 3.3.7). Since the test has low power to detect differences in typical meta-analysis context (few studies) *p* is sometimes set to 0.1 (Higgins, & Green, 2008, sec. 9.5.2) rather than the standard 0.05. The test statistic was used to represent the degree of variability in the treatment effect estimates due to heterogeneity:

$$I^2 = {}^{(Q-df)}/_{Q} \text{ x } 100\%$$

where Q is the χ^2 test-statistic and df is its degrees of freedom (Higgins & Green, 2008). The value of I^2 lies between 0% and 100%, with a value of 0% indicating no observed heterogeneity and larger values show increasing heterogeneity (Higgins, Thompson, Deeks, & Altman, 2003). When there was a sufficient number of studies per meta-analysis we calculated and reported uncertainty intervals for I^2 (Higgins & Thompson, 2002). We found a moderate degree of heterogeneity between studies in some of the meta-analyses, and investigated further when a threshold of 75% was exceeded. In addition we reported the between-studies variance component (τ^2).

3.5.3 Assessment of publication bias

Publication bias may occur because studies that report statistically significant treatment results are more likely to be published. We used two strategies to assess whether publication bias was present. First, our search strategy was designed to uncover any unpublished studies that met inclusion criteria, by searching dissertation databases and grey literature. Second, we examined funnel plots for asymmetry. The funnel plot is constructed by plotting a study's effect size against the standard error of the estimate. In the absence of publication bias one would expect a symmetric graph where the variation in effect size estimates increase with the standard error of the estimate. Although an asymmetric funnel plot does not necessarily imply that publication bias is present, it is an indication that the published research literature may suffer from this type of bias.

3.5.4 Handling of missing data

Not all studies reported sufficient details to allow the calculation of an effect size. When a study reported insufficient data for the calculation of a numeric effect size, we contacted the study authors requesting data. In some cases this allowed us to extract the information needed, but in the majority of cases our attempt to contact authors was unsuccessful[5]. For example, McKay et al. (1999) reported means and standard deviations for drug use but did not report sample size by treatment assignment; only the overall follow-up rate. In this and similar cases we assumed that follow-up attrition was independent of treatment assignment, allowing us to impute the missing data[6]. In addition, as detailed in section 3.4.3.3, when summary statistics were not available, effects sizes were extracted from figures and graphs where possible.

3.5.5 Sensitivity analysis

Sensitivity analysis was carried out with respect to risk of bias items. Meta-analysis was performed excluding studies where the 'sequence generation' item was scored either as High or Unclear, where the 'incomplete data' item was scored at 4, and where the 'other bias' item was scored at 4.

3.5.6 Software used for synthesis

We used Version 5.3 of Review Manager (2014) and Version 13 of Stata (StataCorp, 2013) for data synthesis.

[5] Authors were contacted in September, 2014. Four corresponding authors did not reply to our enquiries, two replied that data had been discarded, and one replied positively and supplied the requested study level information in full detail.
[6] Pigott (2009) describes a number of statistical techniques for dealing with missing data.

4 Results

4.1 RESULTS OF THE SEARCH

We ran the searches in January 2010, October 2011, July 2013, and August 2015, and September 2016.

We searched 15 international and Nordic bibliographic databases, searched for grey literature and hand searched five core journals in September 2016 (see section 3.3 for more information).

The total number of potentially relevant records was 21,974 after excluding duplicates (database search: 17,416; grey literature search: 2,639; hand search and others: 1919).

All 21,974 records were screened based on title and abstract. 428 of these records were retrieved and screened in full text. Thirty-six of the full texts were initially deemed to meet inclusion criteria. Upon closer inspection 23 full texts did not meet inclusion criteria. The primary reason for exclusion of these full texts is listed in section 9.3.

Thirteen reports met the inclusion criteria and data were extracted from these reports.

A total of 10 unique studies, reported in 13 reports, were included in the review. See 'Flow of studies' figure in section 10.2. See section 4.2 for further details on included studies. References to included reports can be found in section 7.1.

4.2 DESCRIPTION OF INCLUDED STUDIES

4.2.1 Study designs

Seven of the included studies were randomised controlled trials (Bisset, 2002; Carroll, Nich, Ball, McCance, & Rounsavile; 1998 Carroll, Nich, Shi, Eagan, & Ball, 2012; Maude-Griffin, Hohenstein, Humfleet, Reilly, Tusel, & Hall, 1998; McKay, Alterman, Cacciola, Rutherford, O'Brien, & Koppenhaver, 1997; Petry, Weinstock, Alessi, Lewis, & Dieckhaus, 2010; Schottenfeld, Moore, & Pantalon, 2011). Two studies had quasi-random allocation of participants to treatment (Higgins, Delaney, Budney, & Bickel, 1991; Wells, Peterson, Gainey, Hawkins, & Catalano, 1994). One quasi-experimental study was included (Gossop, Stewart, & Marsden, 2007).

4.2.2 Location of the studies

All but one of the ten included studies were conducted in the US (Bisset, 2002; Carroll et al., 1998; Carroll et al., 2012; Higgins et al., 1991; Maude-Griffin et al., 1998; McKay et al., 1997; Petry et al., 2010; Schottenfeld et al., 2011; Wells et al., 1994). Gossop et al. (2007) was conducted in the United Kingdom.

All studies were facilitated in outpatient settings.

4.2.3 Participants

As displayed in tables in section 9.2, the number of males and females were approximately equal in three studies (Bisset, 2002; Carroll et al., 2012; Petry et al., 2010). Five studies included a majority of male participants (Carroll et al., 1998; Gossop et al., 2007; Higgins et al., 1991; Maude griffin et al., 1998; Wells et al., 1994). One study included only men (McKay et al., 1997), and another involved only women due to the inclusion criteria of the study that required participants to be female and either pregnant or have custody of a young child (Schottenfeld et al., 2011).

The mean age of the participants varied between 29 and 43 years.

The ten included studies reported the participants' socioeconomic group in five different ways. In two studies, 81% and 62.7% had 12 or more years of education (Bisset, 2002; Higgins et al., 1991). Two studies reported a mean of around 12 years of education (McKay et al., 1997; Petry et al., 2010). Carroll et al. (1998) reported that 23% of participants had some college education, 47% had finished their high school education and 17% had not completed high school. A similar classification was used in Carroll et al. (2012), where the percentages were 34%, 43% and 17%. In two studies, 84% and 95% were either unemployed or did not have a fulltime job (Maude-Griffin et al., 1998; Schottenfeld et al., 2011). In Schottenfeld et al. (2011) it was also reported that 57 % of the participants had at least finished high school or its equivalent. In one study, 68% of the participants had had a full time job the previous 3 years and more than half had worked at least 20 days in the past month (Wells et al., 1994).

Eight studies (Bisset, 2002; Carroll et al., 1998; Carroll et al., 2012; Higgins et al., 1991; Maude-Griffin et al., 1998; McKay et al., 1997; Petry et al., 2010; Schottenfeld et al., 2011) reported participants' mental health conditions either before or when entering treatment. Five studies reported that minimum a third of the participants had some kind of mental disorder (Bisset, 2002; Carroll et al., 1998; Carroll et al., 2012; Maude-Griffin et al., 1998; McKay et al., 1997), including depression, personality disorder, and antisocial personality disorder. One study reported that 24% of the participants met the criteria for current major depression disorder (Schottenfeld et al., 2011). Two studies reported psychiatric problem severity using the Addiction Severity Index Composite Score (ASI) (Higgins et al., 1991; Petry et al., 2010). Two studies did not report the mental health status of participants (Gossop et al., 2007; Wells et al., 1994).

Participants in the included studies had overall a long history of drug use, ranging from 5 to 19 years of drug addiction (Carroll et al., 1998; Gossop et al., 2007; Higgins et al., 1991; McKay et al., 1997; Wells et al., 1994). A number of studies reported drug use among the majority of the

participants within the last 30 days before entering the trial (Bisset, 2002; Carroll et al., 2012; Gossop et al., 2007; Schottenfeld et al., 2011; Wells et al., 1994).

In five studies the participants were predominantly white Caucasian (Bisset, 2002; Carroll et al., 2012; Gossop et al., 2007; Higgins et al., 1991; Wells et al., 1994), in three studies most participants were African American, and two studies reported a combination of ethnic groups (Carroll et al., 1998; Maude-Griffin et al., 1998; McKay et al., 1997; Petry et al., 2010; Schottenfeld et al., 2011).

4.2.4 Interventions

12-step therapy was delivered in a variety of ways across the studies. In four studies, ordinary TSF was delivered (Carroll et al., 1998; Carroll et al., 2012; Maude Griffin et al., 1998; Schottenfeld et al., 2011). The content of the TSF is intended to be consistent with the 12 Steps, but with an importance of steps 1-5 and a disease model of addiction added to the original program. In Bisset (2002), the therapy was delivered as Intensive Twelve Step Facilitation (ITSF). ITSF differs from the ordinary Twelve Step Facilitation (TSF) by offering the participants 48 sessions rather than the 12 session in TSF. 12-step drug counselling was delivered in one study (Higgins et al., 1991). Two studies delivered the therapy as a mix of addiction counselling and 12-step recovery practices (McKay et al., 1997; Petry et al., 2010). In one of these studies, the therapy was delivered as an aftercare program (McKay et al., 1997). In one study the 12-step program was delivered as a self-help group attendance (Gossop et al., 2007).

The duration of the interventions varied between 12 weeks and 6 months. Treatment sessions took place once or twice weekly in all studies.

4.2.5 Control conditions

The control conditions in the included studies was categorised as follows: 12-step therapy was compared to the treatment as usual, another psychosocial intervention or an intervention where 12-step with an add-on was compared to another intervention with an add-on.

In Bisset (2002) participants in one of the two comparison conditions received only methadone and no psychosocial therapy. In one study, one comparison condition was defined as being participants who did not attend any 12-step self-help group meetings (Gossop et al., 2007). The second comparison group in this study consisted of participants who attended self-help sessions less than once a week.

In six of the included studies the comparison conditions were psychosocial interventions without any supplement (Bisset, 2002; Higgins et al., 1991; Maude-Griffin et al., 1998; McKay et al., 1997; Petry et al., 2010; Wells et al., 1994). In one study the comparison condition was Behavioural Therapy (Higgins et al., 1991). Two studies delivered Cognitive Behavioural Therapy to the comparison group (Maude-Griffin et al., 1998; Wells et al., 1994). Acceptance and Commitment Therapy, Contingency Management, and Relapse Prevention interventions were delivered to the comparison group in Bisset (2002), Petry et al. (2010) and McKay et al. (1997), respectively.

In Carroll et al. (1998) participants were divided into five groups. Two groups received CBT and 12-step without add-ons. The remaining three groups all received disulfiram in addition to their

weekly therapy sessions in 12-step, Cognitive Behavioural Therapy and Clinical management. In Carroll et al. (2012) one comparison condition received disulfiram and no psychosocial therapy and the second comparison condition received 12-step and disulfiram. Community reinforcement was delivered to the comparison group, and as add-on the study included contingency and non-contingent yoked voucher control (Schottenfeld et al., 2011).

4.2.6 Treatment fidelity

Therapists with masters or doctoral level training in psychology or with an equivalent degree delivered the treatments in the included studies. One study did not report who delivered the treatment (Gossop et al., 2007).

Seven studies utilised supervision to assess the fidelity of treatment (Bisset, 2002; Carroll et al., 1998; Maude griffin et al., 1998; McKay et al., 1997; Petry et al., 2010; Schottenfeld et al., 2011; Wells et al., 1994).

4.2.7 Measurement of outcomes

An overview of outcomes and how they were measured in each included study can be found in section 11.1.

4.2.7.1 Drug use

The primary outcome for this review was abstinence or reduction in drug use measured by biochemical tests, estimates of drug use, and/or psychometric scales. Eight studies used biochemical tests to verify abstinence (Bisset, 2002; Carroll et al., 2012; Higgins et al., 1991; Maude-Griffin et al., 1998; McKay et al., 1997; Petry et al., 2010; Schottenfeld et al., 2011; Wells et al., 1994).

Six studies used other means of assessing drug use, typically self-report (Carroll et al., 2012; Carroll et al., 1998; Gossop et al., 2007; Maude-Griffin et al., 1 998; McKay et al., 1997; Schottenfeld et al., 2011). In addition, in three studies the ASI-drug subscale was administered (Bisset, 2002; McKay et al., 1997; Schottenfeld et al., 2011).

Continuous outcomes were measured in seven studies (Carroll et al., 1998; Carroll et al., 2012; Gossop et al., 2007; McKay et al., 1997; Petry et al., 2010; Schottenfeld et al., 2011; Wells et al., 1994), whereas dichotomous outcomes were measured in six of the studies (Bisset, 2002; Carroll et al., 1998; Higgins et al., 1991; Maude-Griffin et al., 1998; McKay et al., 1997; Schottenfeld et al., 2011).

4.2.7.2 Secondary outcomes

Two studies measured outcomes in the criminal behaviour domain with the ASI-legal subscale (Bisset, 2002; Carroll et al., 1998). The ASI-legal interview also covers issues related to prostitution.

Three studies used the ASI to assess psychiatric symptoms, social functioning and employment status of the participants (Bisset, 2002; Carroll et al., 1998; McKay et al., 1997). Bisset (2002) also

utilised The Beck Depression Index, and the Symptom Checklist-90-R questionnaires to assess psychological problems and symptoms of psychopathology.

None of the studies measured or reported outcomes related to homelessness.

4.2.8 Time points

Table 1 provides an overview of when measurements were taken in the included studies. The majority of studies took outcome measurements at one or more time points during treatment, except Gossop et al. (2007), Higgins et al. (1991), Petry et al. (2010), and Wells et al. (1994). All studies except Gossop et al. (2007) took outcome measurements at the end of treatment. All studies except Higgins et al. (1991) also took outcome measurements at one or more follow-up points.

Table 1: Time points in included studies

Study	Time points in study post enrolment	Treatment duration
Bisset (2002)	8 weeks 16 weeks 6 months	16 weeks
Carroll et al. (1998)	1 month 2 months 12 weeks 16 weeks 24 weeks 36 weeks 1 year and 12 weeks	12 weeks
Carroll et al. (2012)	1 month 2 months 12 weeks 16 weeks 24 weeks 36 weeks 1 year 1 year and 12 weeks	12 weeks
Gossop et al. (2007)	1 year 2 years 4-5 years	self-help groups
Higgins et al. (1991)	12 weeks	12 weeks
Maude-Griffin et al. (1998)	4 weeks 8 weeks 12 weeks 26 weeks	12 weeks
McKay et al. (1997)	1-6 months 7-12 months 13-18 months 19-24 months	24 weeks
Petry et al. (2010)	24 weeks 6 months	24 weeks

Study	Time points in study post enrolment	Treatment duration
Schottenfeld et al. (2011)	3 months 1-24 weeks 6 months 9 months 12 months	24 weeks
Wells et al. (1994)	12 weeks 36 weeks	12 weeks

4.3 RISK OF BIAS IN INCLUDED STUDIES

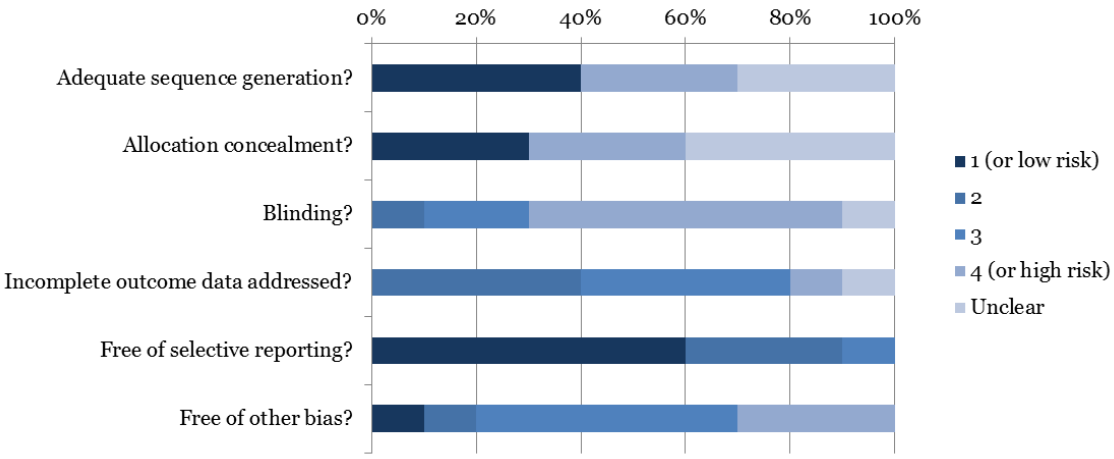
All included studies were assessed for risk of bias. Each study was rated independently by two reviewers in the following 6 domains.

- Selection bias
  - Adequate sequence generation
  - Allocation concealment
- Blinding
- Incomplete outcome data addressed
- Free of selective reporting
- Free of other bias

Some items such as sequence generation were rated on a High risk/Low risk/Unclear risk scale. Other items such as incomplete outcome data were assessed on a scale from 1 to 5, where a score of 5 indicates an unacceptable high risk of bias. Risks of bias tables for each study are presented in section 9.4.1. In all included studies the outcomes were organised into primary outcomes, secondary outcomes and retention within the domains *blinding*, *incomplete outcome data addressed* and *free of selective reporting*. Outcomes within a study received a separate rating if the risk of bias differed between them.

As displayed in Figure 2, four of the included studies received a low risk of bias score for *Adequate sequence generation*. The remaining six studies were either classified as high risk (3) or as unclear risk (3) for sequence generation. Randomisation was performed via a computerised urn-algorithm in studies that were classified with low risk in the adequate sequence generation item. For *allocation concealment*, three studies were rated as low risk of bias, whereas the remaining studies were rated with either high risk (3) or unclear risk (4). The assessment within these two domains indicates that several of the studies are at high risk of some form of selection bias.

Figure 1: Risk of Bias summary for drug use



For *blinding*, none of the included studies received a low risk rating. Six studies were considered as high risk of bias, mainly because of the natural difficulties in blinding therapists and participants.

Eight studies received a medium risk of bias classification (score of 2 or 3) in the *incomplete outcome data addressed* domain. For the domain *free of selective reporting* 6 studies were classified with low risk of bias. The remaining studies were classified with some risk of bias in this domain. One study was judged to be free of *other forms* of bias. Eight studies were given a score of 3 or 4. This was due the lack of correction of the statistical analysis for the group therapy element of the intervention.

The study by Gossop, Stewart, & Marsden (2007) was excluded from meta-analysis following risk of bias assessment. Gossop et al. (2007) used a quasi-experimental design, and accordingly the study was in addition assessed on the confounding item of the risk of bias tool. The study was judged at high risk of bias on the confounding item, leading to its exclusion from quantitative synthesis. Further details on the risk of bias assessment can be found in section 9.5.

4.3.1 Example of risk of bias assessment

Here we present an example of the risk of bias assessment that we conducted for the Higgins et al. (1991) study. Non-random sequence generation was used to assign participants to treatment; 13 participants were allocated to Behavioural therapy and the next 15 participants were allocated to the 12-step group. Because of this allocation, the study was rated with high risk for adequate sequence generation. There was no information in the report regarding allocation concealment. Because the decision to allocate exactly 13 participants to behavioural therapy was not pre-specified, it is likely that this decision was not concealed to study investigators. The study was scored with a high risk rating in allocation concealment.

Neither participants, therapists, nor assessors were blind to treatment allocation, something which is difficult to achieve with these interventions. Because the Behavioural Therapy intervention includes an incentive pay mechanism it is unlikely that staff was blind to the allocation.

For incomplete outcome data/attrition bias, the study was rated with high risk for the primary outcome and with low risk for the outcome of treatment retention. The reason for judging primary outcome at high risk, is due to the rather large, and differential attrition between treatments.

For selective reporting/reporting bias, the study was received a score of 2, indicating a small risk of bias. The reason for the judgement 2 was that reporting on the missing outcome data item was not adequate.

The 12-step intervention was delivered in groups. The statistical analysis in the report did not correct for this design feature. This is assessed in the *other risk of bias* domain.

4.4 EFFECTS OF INTERVENTION

Effect sizes for each study are reported in the online appendix (see section 12) and forest plots for all meta-analyses are reported in Section 11.4.

Prior to calculating effect sizes, outcome measures were recoded such that a higher score was indicative of an improvement in the relevant domain. For example, the ASI is measured on a scale from 0 to 1 where a higher score indicates higher degree of severity. This measure was therefore recoded. Subsequently, effect sizes were constructed such that a positive effect size favoured the 12-step intervention rather than the comparison/control.

Studies used several different scales to measure the same outcome construct. For example, Bisset (2002) measured the primary outcome “drug use” with a urine analysis screen and "relevant portions of ASI" (Hayes et al., 2004, p 673). McKay et al. (1999) used the Timeline Follow Back to construct a measure of percentage of days using cocaine. All effect sizes for drug use were expressed in or transformed to the SMD-family of effect sizes (standardised mean difference). When studies reported multiple effect sizes for drug use, we chose the effect size that relied on a continuous outcome measure.

Treatment retention was also measured in a variety of ways. Some studies measured retention as number of completed sessions (e.g. McKay et al., 1997), others measured retention as treatment completion (e.g. Carroll et al., 1998). Since the majority of studies reported dichotomous outcomes, effect sizes were expressed in or converted to the odds ratio. Studies only contributed one effect size on this outcome construct.

The remaining secondary outcomes were all measured with the relevant subcomponent of the ASI (Addiction Severity Index). Since the ASI is a continuous measure, effect sizes were expressed in the SMD-family of effect sizes. None of the studies contributed multiple effect sizes on these outcome constructs.

A complete list of outcome measures is available in the appendix (section 11.1)

Below we report the results of effect size synthesis. Meta-analysis was conducted separately for two comparisons: 12-step vs. other psychosocial treatment, and 12-step with an add-on (either pharmaceutical add-on such as disulfiram, or another psychosocial add-on such as contingency

management) vs other psychosocial treatment with an identical add-on. Only one study (Bisset, 2002) contributed data to the comparison 12 step vs. no intervention at post treatment and at follow-up. Meta-analysis was therefore not performed, but we report the effect sizes for this study for completeness.

When studies contributed to more than one relevant comparison, these were analysed in separate meta-analyses. Separate meta-analyses were also conducted for each of the outcomes relevant to the review. Finally meta-analysis was conducted separately by time: during intervention, post intervention, and follow-up. The time point division differed from that specified in the protocol. In our protocol we stated that we expected that the included studies would contain interventions of a self-help nature and we anticipated synthesising outcomes in the short-term (less than 6 months after enrolment into treatment), medium-term (6-12 months after enrolment), and long-term (12 months or more after enrolment). However, all studies included in our synthesis were either randomised or quasi-randomised trials. Since most studies reported on the primary outcome during the intervention and at post-intervention, we found it appropriate to follow the research literature in choosing appropriate time intervals[7].

In several studies all or part of the intervention was implemented as group therapy. None corrected for the correlation inherent in these interventions between participants (randomly) assigned to the same group. Since we lacked both plausible information on intra-cluster correlation and on (average) group size, we decided that we could not reliably correct effect sizes; instead the risk of bias assessment for these studies reflected the risk of bias present in the effect estimates for this group of studies.

4.4.1 Comparison: 12-step vs no intervention

4.4.1.1 Primary outcome: drug use

4.4.1.1.1 During treatment

No studies contributed effect sizes for drug use to the comparison of 12-step versus no intervention during treatment.

4.4.1.1.2 Post treatment

One study (Bisset, 2002) contributed effect sizes to the comparison of 12-step versus no intervention at post-treatment. Meta-analysis was therefore not possible. The effect size from Bisset (2002) for drug use is presented in Table 2 below.

[7] Kiluk, Nich, Witkiewitz, Babuscio, & Carroll (2014) show that greater cocaine abstinence during treatment is associated with fewer problems on follow up.

Table 2: Effect size (SMD) for 12-step vs no intervention, outcome: drug use measured post treatment

Outcome	Effect Size (95% CI)
Drug use	0.45 (-0.25:1.14)

Notes: Positive effect size favours 12-step intervention. Effect size for retention: odds ratio.

* $p < .05$, ** $p < .01$

4.4.1.1.3 Follow-up

Only one study contributed a comparison of 12-step versus no intervention (Bisset, 2002), at follow-up, therefore no meta-analysis was possible. Table 3 presents the study level effect size for drug use.

Table 3: Effect size (Hedges' g) for 12-step vs no intervention, outcome: drug use at follow-up (6 months)

Outcome	Effect Size (95% CI)
Drug use	0.91* (0.03:1.78)

Notes. Positive effect size favours 12-step intervention.

* $p < .05$, ** $p < .01$

4.4.1.2 Secondary outcomes

4.4.1.2.1 During treatment

No secondary outcomes were available during treatment for the comparison between 12-step and no intervention.

4.4.1.2.2 Post treatment

Only one study contributed a comparison of 12-step versus no intervention (Bisset, 2002), therefore no meta-analysis was possible. Table 4 presents the study level effect size for the secondary outcomes: retention, criminal behaviour, psychiatric symptoms, social functioning and employment, measured post treatment.

Table 4: Effect sizes (Hedges' g) for 12-step vs no intervention, secondary outcomes measured post treatment

Outcome	Effect Size (95% CI)
Criminal Behaviour	-0.45 (-0.97:0.06)
Psychiatric Symptoms	-0.47 (-0.98:0.04)
Social Functioning	-0.55* (-1.06:-0.03)
Employment	0.20 (-0.31:0.71)
Retention	0.37 (0.14:0.97)

Notes: Positive effect size favours 12-step intervention. Effect size for retention: odds ratio.

* $p < .05$, ** $p < .01$

4.4.1.2.3 Follow-up

Only one study contributed a comparison of 12-step versus no intervention (Bisset, 2002), therefore meta-analysis was not possible. Table 5 presents the study level effect size for the secondary outcomes: retention, criminal behaviour, psychiatric symptoms, social functioning and employment, measured at 6-months follow-up

Table 5: Effect sizes (Hedges' g) for 12-step vs no intervention, secondary outcomes at follow-up (6 months)

Outcome	Effect Size (95% CI)
Criminal Behaviour	-0.11 (-0.65:0.44)
Psychiatric Symptoms	-0.34 (-0.89:0.21)
Social Functioning	0.15 (-0.40:0.69)
Employment	-0.30 (-0.85:0.25)

Notes. Positive effect size favours 12-step intervention.
* $p < .05$, ** $p < .01$

4.4.2 Comparison: 12-step vs other psychosocial interventions

4.4.2.1 Primary outcome: drug use

4.4.2.1.1 During treatment

Five of the seven studies included compared 12-step to a psychosocial intervention contained sufficient information for us to construct an effect size for drug use during treatment (Carroll et al., 1998; Higgins et al., 1991; Maude-Griffin et al., 1998; McKay et al., 1997; Petry et al., 2010). Bisset (2002) and Wells et al. (1994) did not report measures of drug use during treatment.

Table 6 displays the results of meta-analysis for the comparison of 12-step versus other psychosocial intervention for drug use. See Figure 11.4.1.1.1 in Appendix 11.4 for the forest plot. Outcomes were measured during treatment.

Table 6: Meta-analysis, 12-step vs psychosocial intervention, drug use measured during treatment

Outcome	Studies	Participants		Effect Size (95% CI)	Heterogeneity	
		Assigned	In analysis sample		τ^2	I^2 (95% UI)
Drug use	5	502	482	-0.29 (-0.62:0.05)	0,07	57% (0%, 84%)

Notes: Positive effect size favours 12-step intervention. Effect size measure is Hedges' *g*. * $p < .05$, ** $p < .01$ indicates statistical significance of effect size. Uncertainty intervals reported for I^2 where applicable (Higgins & Thompson 2002).

Effect sizes varied between -1.42 (Higgins et al., 1991) and 0.33 (Carroll et al., 1998). All studies except Carroll et al. (1998) included a group element in the intervention, and the estimated standard errors of the effect sizes are therefore likely biased. In the meta-analysis of five studies

(Table 6), there was no statistically significant difference between 12-step and another psychosocial intervention during treatment for the outcome of drug use (Hedges' g = -0.29; 95% CI -0.62 to 0.05). This indicates that 12-step treatment is neither worse nor better than the other psychosocial interventions in reducing drug use. Outcome data were available for 482 out of 502 assigned participants (96%). The I^2 statistic (57%) indicated that there was a moderate degree of heterogeneity present.

4.4.2.1.2 Post treatment

Five of the seven studies that considered the comparison between 12-step and other psychosocial interventions contributed data to the meta-analysis for the outcome of drug use measured post treatment (Bisset, 2002; Carroll et al., 1998; Maude-Griffin et al., 1998; McKay et al., 1997; Wells et al., 1994). Higgins et al. (1991) and Petry et al. (2010) did not measure drug use post treatment.

Table 7 presents results from meta-analysis of the comparison of 12-step intervention versus psychosocial intervention, for drug use, measured post treatment. See Figure 11.4.1.1.2 in Appendix 11.4 for the forest plot.

Table 7: Meta-analysis, 12-step vs psychosocial intervention, drug use measure post treatment

Outcome	Studies	Participants		Effect Size (95% CI)	Heterogeneity	
		Assigned	In analysis sample		τ^2	I^2 (95% UI)
Drug use	5	500	412	-0.03 (-0.24:0.18)	0.00	0% (0%, 73%)

Notes. Positive effect size favours 12-step intervention.

* p < .05, ** p < .01 indicates statistical significance of effect size. Effect size measure is Hedges' g. Uncertainty intervals reported for I^2 where applicable (Higgins & Thompson 2002).

Effect sizes ranged between -0.28 (Maude-Griffin et al., 1998) to 0.25 (Wells et al., 1994). In the meta-analysis of 5 studies (Table 7), there was no statistically significant difference between 12-step and another psychosocial intervention post treatment for the outcome of drug use (Hedges' g = -0.03; 95% CI -0.24 to 0.18). Four of the studies involved group based interventions and data were not corrected for clustering. Data completion rate was 82% on average that is outcome data were available for 82% of participants assigned to treatment. There was no indication of heterogeneity between studies on drug use, but note that the uncertainty interval for I^2 is wide.

4.4.2.1.3 Follow-up

Table 8 presents the results of meta-analyses for the comparison of 12-step treatment versus other psychosocial treatment. Six studies contributed data (Bisset, 2002; Carroll et al., 1998; Maude-Griffin et al., 1998; McKay et al., 1997; Petry et al., 2010; Wells et al., 1994). Outcomes measured 6 months post treatment and 12 months post treatment were synthesised in meta-analysis. Some studies provided follow-up measures at additional time points. For example McKay et al. (1999) provide measures at 18 months follow-up. However 6 and 12 months follow-up times were the time points covered by the most studies, and were chosen for meta-analysis. See Figure 11.4.1.1.3 in Appendix 11.4 for the forest plot. All effect sizes are available in our Online Supplement 1.

Table 8: Meta-analysis, 12-step vs psychosocial intervention, drug use at follow-up

Outcome	Time point (months post treatment)	Studies	Participants		Effect Size (95% CI)	Heterogeneity	
			Assigned	In analysis sample		τ^2	I^2 (95% UI)
Drug use	6	6	670	527	-0.12 (-0.31:0.08)	0.00	0% (0%, 0%)
	12	2	176	146	-0.02 (-0.35:0.3)	0.00	0%

Notes: * $p < .05$, ** $p < .01$ indicates statistical significance of effect size. Uncertainty intervals reported for I^2 where applicable (Higgins & Thompson 2002). Effect size measure is Hedges' g. Positive effect size favours 12-step intervention.

Six studies contributed data to the meta-analysis for drug use at 6 months post treatment (Bisset, 2002; Carroll et al., 1998; Maude-Griffin et al., 1998; McKay et al., 1997; Petry et al., 2010; Wells et al., 1994). Individual effect sizes ranged between -0.29 (Maude-Griffin et al., 1998) and 0.05 (Carroll et al., 1998). In the meta-analysis of 6 studies (Table 8), there was no statistically significant difference between 12-step and another psychosocial intervention at 6 months follow-up for the outcome of drug use (Hedges' g = -0.12; 95% CI -0.31 to 0.08). The average follow-up rate was 79%. In one study (Carroll et al., 1998) treatment was delivered individually. In the remaining studies, treatment was either group therapy alone or a combination of individual and group therapy. Since failing to correct for clustering typically leads to a downward bias in estimated standard errors of the study level effect sizes (leading to too high weights in meta-analysis) the conclusion that there is no statistical difference between 12-step and other psychosocial treatments for the outcome of drug use at 6 months follow-up is likely to be robust to this bias. There was no indication of heterogeneity (I^2 = 0%).

Two studies contributed data to the meta-analysis for the outcome drug use 12 months after treatment end (Carroll et al., 1998; McKay et al., 1997). Study level effect sizes ranged between -0.08 (McKay et al., 1997) and 0.30 (Carroll et al., 1998). In the meta-analysis of these 2 studies (Table 8), there was no statistically significant difference between 12-step and another psychosocial intervention at 12 months follow-up for the outcome of drug use (Hedges' g = -0.02; 95% CI -0.35 to 0.30). The average follow-up rate at 12 months after treatment end was 83%.

4.4.2.2 Secondary outcomes

4.4.2.2.1 During treatment

No studies contributed data for secondary outcomes during treatment.

4.4.2.2.2 Post treatment

Table 9 presents results from the meta-analysis of 12-step vs psychosocial interventions for secondary outcomes measured post treatment. Forest plots for all secondary outcomes can be found in 11.4.1.2 in the Appendix.

Table 9: Meta-analysis, 12-step vs psychosocial intervention, secondary outcomes measured post treatment

Outcome	Studies	Participants		Effect Size (95% CI)	Heterogeneity	
		Assigned	In analysis sample		τ^2	I^2 (95% UI)
Criminal behaviour	2	130	81	-0.31 (-0.76:0.13)	0.00	0%
Psychiatric symptoms	3	262	211	0.08 (-0.19:0.35)	0.00	0% (0%, 84%)
Social functioning	3	262	211	-0.15 (-0.42:0.12)	0.00	0% (0%, 79%)
Employment	3	262	211	0.03 (-0.24:0.3)	0.00	0% (0%, 66%)
Retention¤	6	570	567	0.75* (0.56:1.00)	0.00	1% (0%, 75%)

Notes. ¤ Effect size: odds ratio; Positive effect size favours 12-step intervention (effect size > 1 for retention).

* $p < .05$, ** $p < .01$ indicates statistical significance of effect size. Uncertainty intervals reported for I^2 where applicable (Higgins & Thompson 2002). Effect size measure is Hedges' g.

Two studies contributed data on criminal behaviour (Bisset, 2002; Carroll et al., 1998). Study level effect size estimates were between -0.33 and -0.28. In the meta-analysis of these 2 studies (Table 9), there was no statistically significant difference between 12-step and another psychosocial intervention post-treatment for the outcome of criminal behaviour (Hedges' g = -0.31; 95% CI -0.76 to 0.13). The I^2 statistic indicated a low degree of heterogeneity (0%). An uncertainty interval for I^2 could not be computed due to the small number of studies.

Three studies (Bisset, 2002; Carroll et al., 1998; McKay et al., 1997) contributed data on psychiatric symptoms, employment and social functioning. The data completion rate for these three outcomes was 81%. Two of the studies used individual randomisation to group therapy.

Effect size estimates for psychiatric symptoms ranged between -0.13 and 0.20. In the meta-analysis of 3 studies (Table 9), there was no statistically significant difference between 12-step and another psychological intervention post-treatment for the outcome of psychiatric symptoms (Hedges' g = 0.08; 95% CI -0.19 to 0.35). The I^2 statistic indicated a low degree of heterogeneity (0%), but note that the uncertainty interval for I^2 is wide.

Study level effect sizes for social functioning ranged between -0.2 and 0.22. In the meta-analysis of 3 studies (table 9), there was no statistically significant difference between 12-step and another psychosocial intervention post-treatment for the outcome of social functioning (Hedges' g = -0.15; 95% CI -0.42 to 0.12). The I^2 statistic indicated a low degree of heterogeneity (0%), but note that the uncertainty interval for I^2 is wide.

Employment effect size estimates ranged between -0.03 and 0.32. In the meta-analysis of 3 studies (table 9), there was no statistically significant difference between 12-step and another psychosocial intervention post-treatment for the outcome of employment (Hedges' g = 0.03; 95% CI -0.24 to

0.3). The I^2 statistic indicated a low degree of heterogeneity (0%), but note that the uncertainty interval for I^2 is wide.

Six studies (Bisset, 2002; Carroll et al., 1998; Higgins et al., 1991; McKay et al., 1997; Petry et al., 2010; Wells et al., 1994) contributed data to the meta-analysis for the outcome treatment retention. Since most studies reported retention as a binary variable, the odds ratio effect size was used for synthesis. The data completion rates were above 99%. Five out of the six studies involved group based treatments. The summary effect size was statistically significant at the 5% level (OR = 0.75; 95% CI 0.56 to 1.00). 12-step interventions thus appear to do worse at retaining participants in treatment than the other psychosocial interventions. Five of the studies did not correct for the fact that outcome analysis was conducted at the individual level, but allocation involved individual randomisation to group therapy. As such there is a risk that the estimated standard error of the effect size is too small. Therefore the conclusion that 12-step interventions appear worse at retaining participants in treatment than other psychosocial interventions should be drawn cautiously. The I^2 statistic indicated a low degree of heterogeneity between studies (I^2 = 1%), but note that the uncertainty interval for I^2 is wide.

4.4.2.2.3 Follow-up

For the secondary outcomes (psychiatric symptoms, employment, social functioning) a total of three studies contributed data at 6 months follow-up (Bisset 2002; Carroll et al., 1998; McKay et al., 1997); only two of the three (Carroll et al., 1997; McKay et al. 1997) contributed data at 12 months follow-up. Average follow-up rate was 73% at 6 months follow-up, and 80% at 12 months follow-up. Carroll et al. (1998) and Bisset (2002) contributed data for the outcome of criminal behaviour at 6 months follow-up (55% follow-up rate). Only Carroll et al. (1998) provided follow-up data for this outcome measure at 12 months follow-up. Accordingly meta-analysis was not possible for this time point but we have reported the study level effect size in Table 10. Forest plots for secondary outcomes at follow-up can be found 11.4.1.2.2 in the Appendix.

Table 10: Meta-analysis, 12-step vs psychosocial intervention, secondary outcomes at follow-up

Outcome	Time point (months post treatment)	Studies	Participants		Effect Size (95% CI)	Heterogeneity	
			Assigned	In analysis sample		τ^2	I^2 (95% UI)
Criminal behaviour	6	2	130	71	-0.17 (-0.78:0.45)	0.07	35%
	12	1	34	19	-0.32 (-1.29:0.66)	NA	NA
Psychiatric symptoms	6	3	262	192	-0.06 (-0.38:0.25)	0.01	10% (0%, 91%)
	12	2	176	140	0.02 (-0.32:0.35)	0.00	0%
Social functioning	6	3	262	192	0.02 (-0.26:0.31)	0.00	0% (0%, 0%)
	12	2	176	141	-0.06 (-0.4:0.27)	0.00	0%
Employment	6	3	262	191	-0.07 (-0.36:0.22)	0.00	0% (0%, 90%)
	12	2	176	140	-0.18 (-0.52:0.15)	0.00	0%

Notes: * $p < .05$, ** $p < .01$ indicates statistical significance of effect size. Uncertainty intervals reported for I^2 where applicable (Higgins & Thompson 2002). Positive effect size favours 12-step intervention. Effect size measure is Hedges' g.

For all outcomes and time points the estimated summary effect sizes were not statistically significant (range: -0.18 to 0.02). Between study heterogeneity ranged from low to moderate (I^2 between 0% and 35%).

4.4.3 Comparison: 12-step with add-on vs other psychosocial interventions with add-on

4.4.3.1 Primary outcome: drug use

4.4.3.1.1 During treatment

Table 11 presents meta-analytical results for the comparison of 12-step intervention with an add-on treatment (either pharmaceutical or psychosocial) against a psychosocial intervention with the same add-on for drug use. Outcome measurements were taken during treatment. Three studies contributed effect sizes to the meta-analysis (Carroll et al., 1998; Carroll et al., 2012; Schottenfeld et al., 2011). The comparison in Carroll et al. (1998) was between 12-step and either cognitive behavioural therapy (CBT) or clinical management (ClM). Disulfiram (Dis) was used as an add-on. In Carroll et al. (2012) the comparison was between 12-step and clinical treatment-as-usual (TAU). Either disulfiram (Dis) or placebo (Pla) was used as an add-on, which gave rise to two relevant comparisons: 12+Dis v TAU+Dis and 12+Pla v TAU+Pla. Schottenfeld et al. (2011) examined the comparison between 12-step and community reinforcement approach (CRA) interacted with either

contingency management (CM) or voucher control (VC). Forest plots are presented in 11.4.2.1 in the Appendix.

Table 11: Meta-analysis, 12-step with add-on vs psychosocial intervention with add-on, drug use measured during treatment

Outcome	Comparison		Studies	Participants		Effect Size (95% CI)	Heterogeneity	
	Main Comparison: 12- step vs.	Add-on		Assigned	In analysis sample		τ^2	I^2 (95% UI)
Drug use	CBT, CRA, TAU	Dis, CM	3[a]	183	154	0.15 (-0.17:0.46)	0.00	0% (0%, 47%)
		Dis, CM, Pla	3[b]	177	152	0.3 (-0.15:0.75)	0.07	45% (0%, 83%)
		Dis, VC	3[c]	182	152	0.11 (-0.21:0.42)	0.00	0% (0%, 67%)
		Dis, VC, Pla	3[d]	176	150	0.26 (-0.23:0.76)	0.10	54% (0%, 86%)
	ClM, CRA, TAU	Dis, CM	3[e]	184	152	0.18 (-0.14:0.50)	0.00	0% (0%, 39%)
		Dis, CM, Pla	3[f]	178	150	0.35 (-0.07:0.77)	0.05	37% (0, 80%)
		Dis, VC	3[g]	183	150	0.14 (-0.18:0.46)	0.00	0% (0%, 70%)
		Dis, VC, Pla	3[h]	177	148	0.32 (-0.16:0.80)	0.09	50% (0%, 85%)

Notes. Main comparison: CBT = Cognitive Behavioural Therapy, CRA = Community Reinforcement Approach, TAU = Treatment as Usual (Standard Counseling), ClM =Clinical Management. Add-ons: Dis = Disulfiram, CM = Contingency Management, Pla = Placebo, VC = Voucher Control. Positive effect size favours 12-step intervention. Effect size measure is Hedges' *g*.

* $p < .05$, ** $p < .01$ indicates statistical significance of effect size. Uncertainty intervals reported for I^2 where applicable (Higgins & Thompson 2002).

[a] Carroll 1998: 12+Dis v CBT+Dis, Schottenfeld 2011: 12+CM v CRA+CM, Carroll 2012: 12+Dis v TAU+Dis; [b] Carroll 1998: 12+Dis v CBT+Dis, Schottenfeld 2011: 12+CM v CRA+CM, Carroll 2012: 12+Pla v TAU+Pla; [c] Schottenfeld 2011: 12+VC v CRA+VC, Carroll 1998: 12+Dis v CBT+Dis, Carroll 2012: 12+Dis v TAU+Dis; [d] Schottenfeld 2011: 12+VC v CRA+VC, Carroll 1998: 12+Dis v CBT+Dis, Carroll 2012: 12+Pla v TAU+Pla; [e] Schottenfeld 2011: 12+CM v CRA+CM, Carroll 1998: 12+Dis v ClM+Dis, Carroll 2012: 12+Dis v TAU+Dis; [f] Schottenfeld 2011: 12+CM v CRA+CM, Carroll 1998: 12+Dis v ClM+Dis, Carroll 2012: 12+Pla v TAU+Pla; [g] Schottenfeld 2011: 12+VC v CRA+VC, Carroll 1998: 12+Dis v ClM+Dis, Carroll 2012: 12+Dis v TAU+Dis; [h] Schottenfeld 2011: 12+VC v CRA+VC, Carroll 1998: 12+Dis v ClM+Dis, Carroll 2012: 12+Pla v TAU+Pla.

Average effect sizes estimates ranged from 0.11 to 0.35. None of the summary effect size estimates for the outcome of drug use were statistically significant at the 5% level. Data completion rates varied between 82-86%. All of the studies were randomised trials where treatment was delivered individually. Across the different analyses, the degree of heterogeneity between studies (I^2) varied from low to moderate (0-54%).

4.4.3.1.2 Post treatment

In Table 12 we report meta-analytical results for the comparison 12-step with an add-on versus other psychosocial treatment with an identical add-on for the primary outcome of drug use. Forest plots are presented in 11.4.2.1.2 in the Appendix.

Table 12: Meta-analysis, 12-step with add-on vs psychosocial intervention with add-on, drug use measured post treatment

Outcome	Comparison		Studies	Participants		Effect Size (95% CI)	Heterogeneity	
	Main comparison: 12 step vs.	Add-on		Allocated	In analysis sample		τ^2	I^2 (95% UI)
Drug use	CBT, TAU	Dis	2[a]	108	94	0.07 (-0.64:0.78)	0.18	66%
		Dis, Pla	2[b]	104	89	-0.01 (-0.56:0.54)	0.07	41%
	ClM, TAU	Dis	2[c]	111	96	0.29 (-0.11:0.70)	0.00	0%
		Dis, Pla	2[d]	105	91	0.20 (-0.22:0.61)	0.00	0%

Notes. Main comparison: CBT = Cognitive Behavioural Therapy, CRA = Community Reinforcement Approach, TAU = Treatment as Usual (Standard Counseling), ClM =Clinical Management. Add-ons: Dis = Disulfiram, CM = Contingency Management, Pla = Placebo, VC = Voucher Control.

* $p < .05$, ** $p < .01$ indicates statistical significance of effect size. Effect size measure is Hedges' g. Uncertainty intervals reported for I^2 where applicable (Higgins & Thompson 2002).; Positive effect size (odds ratio above 1) favours 12-step intervention; ¤ odds ratio;

[a] Carroll 1998: 12+Dis v CBT+Dis, Carroll 2012: 12+Dis v TAU+Dis; [b] Carroll 1998: 12+Dis v CBT+Dis, Carroll 2012: 12+Pla v TAU+Pla; [c] Carroll 1998: 12+Dis v ClM+Dis. Carroll 2012: 12+Dis v TAU+Dis; [d] Carroll 1998: 12+Dis v ClM+Dis, Carroll 2012: 12+Pla v TAU+Pla

Two studies contributed data to the meta-analysis on drug use (Carroll et al., 1998; Carroll et al., 2012). 12-step was compared to either cognitive behavioural therapy (CBT), clinical management (ClM) or treatment-as-usual (TAU). Disulfiram (Dis) and placebo (Pla) were add-ons. The main treatments were delivered individually. Individual effect sizes ranged between -0.32 (Carroll et., 1998; 12-step and disulfiram versus CBT and disulfiram) to 0.41 (Carroll et al., 2012; 12-step and disulfiram in addition to treatment-as-usual versus treatment-as-usual and disulfiram). None of the estimated summary effect sizes (range: -0.01 to 0.29) for drug use at post-treatment were statistically significant. Data completion rates were 86-87%. Heterogeneity was moderate (I^2 ranging from 0 to 66%). An uncertainty interval could not be computed due to the low number of studies.

4.4.3.1.3 Follow-up

Table 13 below presents results from meta-analyses where 12-step with an add-on was compared against other psychosocial intervention with the same add-on for drug use. Only two studies (Carroll et al., 1998; Carroll et al., 2012) contributed to the meta-analysis at 6 and 12 months post

treatment[8]. 12-step was compared to either cognitive behavioural therapy (CBT), clinical management (ClM), or treatment-as-usual (TAU). Disulfiram (Dis) and placebo (Pla) were add-ons. Forest plots are presented in 11.4.2.1.3 in the Appendix.

Table 13: Meta-analysis, 12-step with add-on vs psychosocial intervention with add-on, drug use at follow-up

Outcome	Comparison		Studies	Time point (months post treat-ment)	Participants		Effect Size (95% CI)	Heterogeneity	
	Main comparison: 12 step vs.	Add on			Assigned	In analysis sample		τ^2	I^2
Drug use	CBT, TAU	Dis	2[a]	6	110	87	0.17 (-0.52:0.86)	0.15	60%
				12		84	0.08 (-0.56:0.71)	0.11	50%
		Dis, Pla	2[b]	6	104	84	0.14 (-0.50:0.79)	0.12	53%
				12		82	0.03 (-0.49:0.54)	0.04	26%
	ClM, TAU	Dis	2[c]	6	111	90	0.48* (0.06:0.90)	0,00	0%
				12		84	0.27 (-0.16:0.71)	0.00	0%
		Dis, Pla	2[d]	6	105	87	0.45* (0.03:0.88)	0,00	0%
				12		82	0.20 (-0.23:0.64)	0,00	0%

Notes: Main comparison: CBT = Cognitive Behavioural Therapy, TAU = Treatment as Usual (Standard Counseling), ClM = Clinical Management. Add-ons: Dis = Disulfiram, Pla = Placebo. ; Positive effect size favours 12-step intervention;

* $p < .05$, ** $p < .01$ indicates statistical significance of effect size. Effect size measure is Hedges' g.

[a] Carroll 1998: 12+Dis v CBT+Dis, Carroll 2012: 12+Dis v TAU+Dis; [b] Carroll 1998: 12+Dis v CBT+Dis, Carroll 2012: 12+Pla v TAU+Pla; [c] Carroll 1998: 12+Dis v ClM+Dis, Carroll 2012: 12+Dis v TAU+Dis; [d] Carroll 1998: 12+Dis v ClM+Dis, Carroll 2012: 12+Pla v TAU+Pla.

Follow-up rates were between 76% and 83%. Two of the summary effect sizes for drug use at 6 months follow-up were statistically significant at the 5% level. Both summary effect sizes favoured the 12-step intervention. The estimated summary effect size for drug use at 6 months for 12-step versus TAU or Clinical Management with disulfiram add-on was statistically significant (Hedges' *g* = 0.48; 95% CI 0.06 to 0.90). The meta-analysis was based on an analysis sample of 90 participants from 2 studies (Caroll et al., 1998; Caroll et al., 2012). The estimated summary effect size for drug use at 6 months for 12-step versus TAU (with placebo add-on) or Clinical Management (with disulfiram add-on) was statistically significant (Hedges' *g* = 0.45; 95% CI 0.03 to 0.88). The meta-analysis was based on an analysis sample of 87 participants from 2 studies (Caroll et al., 1998; Caroll et al., 2012). The results of the remaining two analyses for drug use at 6 months follow-up were not statistically significant. These meta-analyses were based on 2 studies

[8] Since only two studies were available for synthesis calculating uncertainty intervals for I^2 was not possible.

(Caroll et al., 1998; Caroll et al., 2012). In one analysis 87 participants contributed data where 12-step was compared to CBT and TAU (with disulfiram as add-on). In the other analysis, 84 participants contributed data to analysis where 12-step was compared to CBT (with disulfiram as add-on) and TAU (with placebo as add-on).

Analyses of the outcome of drugs use conducted 12 months after end of treatment did not reveal statistically significant differences between 12-step with an add-on treatment and the alternative intervention(s) with add-on treatment.

Conclusions must be drawn cautiously. Both studies contributing data were from the same team of investigators and dependencies between study level effect sizes may contribute to exaggerated *p*-values.

Across all the meta-analyses reported in Table 13, there was indication of a low to medium degree of study heterogeneity (I^2 statistic ranged from 0% to 60%).

4.4.3.2 Secondary outcomes

4.4.3.2.1 During treatment

No studies contributed data on secondary outcomes during treatment.

4.4.3.2.2 Post treatment

Three studies (Carroll et al., 1998; Carroll et al., 2012; Schottenfeld et al., 2011) contributed data on treatment retention (Table 14). Forest plots are presented in 11.4.2.2.1 in the Appendix.

Table 14: Meta-analysis, 12-step with add-on vs psychosocial intervention with add-on, treatment retention measured post treatment

Outcome	Comparison		Studies	Participants		Effect Size (95% CI)	Heterogeneity	
	Main comparison: 12 step vs.	Add-on		Allocated	In analysis sample		τ^2	I^2 (95% UI)
Retention¤	CBT, CRA, TAU	Dis, CM	3[e]	183	183	0.70 (0.38:1.29)	0.00	0% (0%, 64%)
		Dis, CM, Pla	3[f]	177	177	0.74 (0.39:1:39)	0.00	0% (0%, 31%)
		Dis, VC	3[g]	182	182	0.73 (0.40:1.33)	0.00	0% (0%, 67%)
		Dis, VC, Pla	3[h]	176	176	0.77 (0.41:1.44)	0.00	0% (0%, 35%)
	CRA, ClM, TAU	Dis, CM	3[i]	184	184	0.78 (0.42:1.44)	0.00	0% (0%, 86%)
		Dis, CM, Pla	3[j]	178	178	0.83 (0.44:1.56)	0.00	0% (0%, 81%)
		Dis, VC	3[k]	183	183	0.81 (0.44:1.49)	0.00	0% (0%, 86%)
		Dis, VC, Pla	3[l]	177	177	0.87 (0.46:1.62)	0.00	0% (0%, 79%)

Notes. ¤ odds ratio; Main comparison: CBT = Cognitive Behavioural Therapy, CRA = Community Reinforcement Approach, TAU = Treatment as Usual (Standard Counseling), ClM =Clinical Management. Add-ons: Dis = Disulfiram, CM = Contingency Management, Pla = Placebo, VC = Voucher Control.
* $p < .05$, ** $p < .01$ indicates statistical significance of effect size. Effect size measure is Hedges' g. Uncertainty intervals reported for I^2 where applicable (Higgins & Thompson 2002).; Positive effect size (odds ratio above 1) favours 12-step intervention;
[e] Carroll 1998: 12+Dis v CBT+Dis, Schottenfeld 2011: 12+CM v CRA+CM, Carroll 2012: 12+Dis v TAU+Dis; [f] Carroll 1998: 12+Dis v CBT+Dis, Schottenfeld 2011: 12+CM v CRA+CM, Carroll 2012: 12+Pla v TAU+Pla; [g] Schottenfeld 2011: 12+VC v CRA+VC, Carroll 1998: 12+Dis v CBT+Dis, Carroll 2012: 12+Dis v TAU+Dis; [h] Schottenfeld 2011: 12+VC v CRA+VC, Carroll 1998: 12+Dis v CBT+Dis, Carroll 2012: 12+Pla v TAU+Pla; [i] Schottenfeld 2011: 12+CM v CRA+CM, Carroll 1998: 12+Dis v ClM+Dis, Carroll 2012: 12+Dis v TAU+Dis; [j] Schottenfeld 2011: 12+CM v CRA+CM, Carroll 1998: 12+Dis v ClM+Dis, Carroll 2012: 12+Pla v TAU+Pla; [k] Schottenfeld 2011: 12+VC v CRA+VC, Carroll 1998: 12+Dis v ClM+Dis, Carroll 2012: 12+Dis v TAU+Dis; [l] Schottenfeld 2011: 12+VC v CRA+VC, Carroll 1998: 12+Dis v ClM+Dis, Carroll 2012: 12+Pla v TAU+Pla.

Study effect sizes (odds ratios) ranged between 0.50 (Carroll et al., 2012; 12-step and disulfiram in addition to treatment-as-usual versus disulfiram and treatment-as-usual) and 1.33 (Carroll et al, 1998; 12-step and disulfiram versus clinical management and disulfiram). Estimated summary effect sizes ranged between 0.73 and 0.87. None of the estimated summary effect sizes for treatment retention were statistically significant. The I^2 statistic indicated a low degree of heterogeneity (0%).

4.4.3.2.3 Follow-up

No studies contributed data to secondary outcomes at follow-up.

4.5 SENSITIVITY ANALYSIS

We conducted sensitivity analysis for drug use and treatment retention for the comparison 12-step against other psychosocial intervention. Separate sensitivity analyses were conducted (when possible) for drug use during treatment, post treatment, and six months after end of treatment. These were the only outcomes where the number of studies contributing data was sufficient for sensitivity analysis to be meaningful. We examined the sensitivity of the results to the risk of bias items: sequence generation, incomplete data, and other bias. These analyses did not uncover any significant changes in the point estimates of the summary effect size. Given the small number of studies available, the power to detect such changes was also low (see section 11.5).

4.6 PUBLICATION BIAS

In attempt to investigate whether the available data might indicate publication bias, we constructed funnel plots for two outcomes: drug use (during treatment, post treatment, and at 6 months follow-up), and treatment retention. Inspection of the funnel plots did not indicate the presence of publication bias (see section 11.6). However, the power to detect the presence of publication bias is limited by the small number of available studies (Higgins, & Green, 2008, chapter 10.4.3.1).

5 Discussion

5.1 SUMMARY OF MAIN RESULTS

The main objective of this review was to synthesise the effects of 12-step programs aimed at reducing drug use of individuals who use illicit drugs. After an extensive systematic search we found a total of ten studies that met the inclusion criteria for the review. In total we found ten studies meeting our inclusion criteria. Seven of the included studies used a RCT design, two studies used a QRCT design, and one study used a QES design. In seven studies, treatment was partially or fully delivered in group therapy sessions. One study, assessed at high risk of bias, was excluded from meta-analysis. Thus nine studies with a total of 1,071 participants contributed data to the analyses.

In the meta-analyses we compared a) 12-step intervention to other psychosocial intervention(s), and b) 12-step with an add-on treatment versus other psychosocial intervention(s) with an (identical) add-on treatment. Seven studies contributed data to the comparison of 12-step intervention to alternative psychosocial interventions during treatment, at treatment end, and at follow-up (Bisset, 2002; Carroll et al., 1998; Higgins et. al, 1991; Maude-Griffin et al., 1998; McKay et al., 1997; Petry et al. 2010; Wells et al., 1994). The seven studies did not all contribute data to all time points. Three studies contributed data to the comparison between 12-step with an add-on and other psychosocial interventions with the same add-on during treatment, at treatment end, and at follow-up (Carroll et al., 1998; Carroll et al., 2012; Schottenfeld et al., 2011). The primary outcome of interest was reduction in drug use or abstinence from drugs. The secondary outcomes considered were criminal behaviour, prostitution, psychiatric symptoms, social functioning, employment status, homelessness, and retention in treatment. We used random-effects meta-analysis to synthesise effects for each comparison, outcome, and time point (during treatment, post treatment, and at follow-up).

The main bulk of evidence refers to the comparison of 12-step interventions versus other psychosocial interventions. At each time point where a separate meta-analysis was conducted, the estimated summary effect size for drug use was not statistically significant, and therefore did not favour either 12-step or the alternative set of interventions. An additional note of caution stems from the fact that six out of seven studies (Bisset, 2002; Higgins et. al, 1991; Maude-Griffin et al., 1998; McKay et al., 1997; Petry et al. 2010; Wells et al., 1994) did not correct for the design-effect that resulted from individual participants being randomised to group therapy (IRGT-design).

The evidence from comparing 12-step with an add-on treatment against psychosocial intervention with an (identical) add-on treatment for the outcome of drug use were cautiously in favour of 12-step at 6 months follow-up. The evidence builds on data contributed from at most three studies (Carroll et al., 1998; Caroll et al., 2012; Schottenfeld et al., 2011), two of which have the same principal investigator. Summary effect sizes favouring 12-step reached statistical significance at 6 months follow-up, but not at 12 months follow-up. Because multiple testing is involved the tests are unlikely to have nominal size. One of the three studies utilised an IRGT-design (Carroll et al., 2012). We found no statistically significant differences for other outcomes or time points.

All included studies measured and reported on treatment retention. Six studies contributed data for retention to the meta-analysis of the comparison between 12-step and other psychosocial interventions. The summary effect size (OR) was estimated as 0.75 (CI 95%: 0.56 to 1.00) and this effect was significant at the 5% level. This suggests that the alternative set of interventions had more success retaining participants. However five of the studies employed an IRGT-design, which likely exaggerates the precision of the estimated effect sizes. Other comparisons revealed no statistically significant summary effect sizes.

Three studies (Bisset, 2002; Carroll et al., 1998, McKay et al., 1997) reported on a secondary outcome: criminal behaviour, psychiatric symptoms, social function and employment. Meta-analysis for these outcomes did not reveal any statistically significant summary effect sizes. Our secondary outcomes included prostitution and homelessness. None of the studies reported directly on these outcomes, therefore we were not able to present any evidence about the effectiveness of the intervention of these two secondary outcomes, although the legal subcomponent of the ASI, administered in three studies (Bisset, 2002; Carroll et al., 1998; McKay et al., 1997), contains indirect information on prostitution.

5.2 OVERALL COMPLETENESS AND APPLICABILITY OF EVIDENCE

We performed a comprehensive electronic database search, combined with grey literature searching, and hand searching of key journals. In addition, experts in the field were consulted regarding potentially missing studies from our list of included studies. No studies are awaiting classification, nor have we been made aware of any on-going studies. All citations were screened by two independent screeners.

Contrary to our expectations, our search uncovered only one quasi-experimental study meeting the inclusion criteria (Gossop et al., 2007). Gossop et al. (2007) was not included in any meta-analysis because it was deemed to entail too high of a risk of bias. In the context of 12-step, it is surprising that we only located one quasi-experimental study report in our search. Narcotics Anonymous has over 58,000 weekly meetings worldwide (Narcotics Anonymous, 2012). This is indicative of a discrepancy between practice and evidence. In this sense the results of this review regarding effects of 12-step in other settings must be viewed cautiously. This lack of evidence is an important caveat in terms of understanding how and whether the results are applicable to 12-step interventions in self-help settings.

Our main result is that 12-step interventions are, given the available evidence, neither better nor worse than competing interventions. In this situation the cost of the intervention is an important parameter for a decision maker. With a fixed budget and two equally effective interventions, an agency is able to treat more patients if the budget is spent on the less expensive intervention. From a cost perspective, self-help groups with their emphasis on peers, role modelling and social support, appear particularly attractive. We did not study this issue systematically, and included studies did not report on the cost of implementation.

5.3 QUALITY OF THE EVIDENCE

The results of this review should be seen in the context of the weight of evidence. Ten unique studies met inclusion criteria, and not all of these studies contributed data to the same comparisons.

The target population for studies included in this review is notoriously difficult to treat, retain, and follow-up. This results in attrition despite the best efforts of primary researchers to follow-up the complete study sample. As a result, none of the studies were able to perform a true intention-to-treat analysis. For the outcome of drug use for the comparison of 12-step to other psychosocial interventions during treatment the average data completion rate (percent of participants assigned where outcome data was available for analysis) was 96%. The data completion rate for drug use measured post-treatment and at follow-up was lower, in both cases around 80%. Most studies imputed missing values by, for example, assuming that missing urine samples were positive. Other studies limited attention to the part of the sample they were able to follow-up with. Some of the more recent studies used more advanced statistical techniques such as regression analysis that allow for missing data.

Perhaps the biggest concern is the level of analysis. Only two of the included studies carried out statistical analysis at the same level that randomisation was done. The remaining studies employed an IRGT-design (Pals et al., 2008), where participants were individually allocated to treatment, but the treatment itself was administered either fully or partially as group therapy. In these studies, the level of statistical analysis was not conducted at the group level, but rather at the individual level, ignoring the dependency between participants assigned to the same group. If estimates of within-cluster variance and group size are available, a design effect (Kish, 1965) can be calculated and used to correct study level standard errors of effect size estimates. Because studies neither reported intra-cluster correlations, nor group sizes, and no credible auxiliary evidence could be brought to bear, we did not correct effect sizes for this design feature. Instead, studies were scored in the Other Bias item in the risk of bias assessment. Regardless, this feature of the evidence warrants caution. When the statistical analysis does not account for clustering there is a risk that the estimated standard errors of effect sizes are too small.

We did not find any clear indication of reporting bias. Authors generally reported all results relevant to the review when they had stated that outcomes were measured. On the other hand, this does not imply that reporting bias is not present. For example, only two of the seven RCT studies reported having registered a protocol (Carroll et al., 2012; Schottenfeld et al., 2011). In the cases

where investigators reported having measured additional outcomes relevant to the review, and when these outcomes were not available in the published report, we were able to contact study investigators and obtain the data.

The meta-analyses did not suggest a high degree of statistical heterogeneity. We performed a sensitivity analysis on the basis of the study level risk of bias assessments. These analyses suggested that conclusions were robust with respect to excluding studies with a high risk of bias on, respectively, sequence generation, incomplete data, and other bias. We did not find evidence of publication bias, but the number of studies examined was also small, implying low statistical power to detect bias.

5.4 LIMITATIONS AND POTENTIAL BIASES IN THE REVIEW PROCESS

Our search strategy did not include any language restrictions. Even so it is noteworthy that all but one of the included studies were conducted in the US. Our literature search was also conducted on non-English electronic databases, and the grey literature search was directed toward non-English sources. Thus we have no particular reason to believe that any language bias has been introduced.

In our protocol we did not state that we would assess adverse events. This is a potential limitation, in as far as 12-step interventions might have more or fewer adverse events compared to other interventions. Adverse events were not systematically reported in study reports, and 12-step programs were not reported as leading to additional adverse events compared to other interventions.

The majority of studies employed a study design where participants were individually randomised to group treatment. Such procedures introduce issues with the unit of analysis, because studies conducted analysis at the individual level. This was coded in the “Other bias” item of the risk of bias tool.

We did not include cost as an outcome of interest and none of the included studies supplied data on the cost of implementation.

5.5 AGREEMENTS AND DISAGREEMENTS WITH OTHER STUDIES OR REVIEWS

To the best of our knowledge there is no review that compares 12-step interventions to no intervention or to other interventions for illicit drug users. There are reviews that compare specific drug treatment programs to a condition such as methadone maintenance (Marsch, 1998) and contingency management (Griffith, Rowan-Szal, Roark, & Simpson, 2000) that could be viewed as ‘no intervention’. They found that the specific drug treatment programs were effective compared to no intervention. Prendergast, Podus, Chang, and Urada (2002) compared drug abuse treatments to a no/minimal treatment and to treatment as usual and found a statistically significant summary effect size around 0.30.

Ferri, Amato, and Davoli (2006) reviewed AA and other 12-step programs (TSF) for alcohol dependence and included eight trials. They tentatively concluded that AA/TSF is not more effective than competing interventions; a finding this review echoes, on a different population.

6 Authors' conclusions

6.1 IMPLICATIONS FOR PRACTICE AND POLICY

This review set out to compare 12-step programs (TSF and self-help) for illicit drug dependence to no/minimal intervention and to competing interventions.

The review located only one study that compared 12-step to minimal intervention (Bisset, 2002). In this context it is important to emphasise that other reviews have demonstrated that drug abuse treatment is effective in reducing drug use compared to no/minimal intervention (see for example, Prendergast et al. 2002 and the references cited therein).

The main bulk of evidence of this review compares 12-step approaches to other competing psychosocial approaches to drug treatment, such as contingency management, CBT, BT, and relapse prevention. We also found a smaller evidence base that compared 12-step with an add-on to other psychosocial interventions with an add-on.

The primary outcome of drug use was assessed during treatment, post treatment, and at 6 and 12 months follow-up. We found no statistically significant summary effect size estimates for drug use during treatment and post treatment. At 6 months follow-up we found that 12-step with an add-on compared to other interventions with an identical add-on led to lower drug use (Hedges' g =0.48, 95% CI: 0.06 to 0.90, and g=0.45, 95% CI: 0.03 to 0.88). Other comparisons at 6 months follow-up for drug use were not statistically significant. No statistically significant differences between treatments for drug use were found at 12-months follow-up.

There are a number of reasons to interpret results cautiously. First, the population is hard to retain in treatment and even harder to follow-up. This increases the uncertainty with which effects of interventions are measured. For example very few studies report intention-to-treat (ITT) estimates. In the context of voluntary participation in programs an ITT estimate of effectiveness is most relevant when selecting between competing programs. Second, most drug abuse treatments are implemented in group therapy settings. None of the studies who delivered treatment in group therapy settings adequately dealt with this design feature in their statistical analysis. The implication is that the precision of effect size estimates from individual studies will tend to be exaggerated.

Based on these results our interpretation of the evidence is that 12-step programs appear to be as effective as competing drug abuse treatments. The evidence is inconclusive about whether one alternative program is better than another. This implies that local context and experience should take precedence when choosing between competing treatments.

The participants in the included studies were both males and females. Across studies the average age of participants ranged from 28 to 43 years on average. Participants were predominantly cocaine dependent, and in two studies participants were also on methadone maintenance (Bisset, 2002; Carroll et al., 2012). Two studies considered participants that were particularly vulnerable who in addition to drug dependence were either: HIV positive (Petry et al., 2010) or pregnant/had young dependants (Schottenfeld et al., 2011). This review offers limited evidence about the effects of treatment for other types of illicit drug dependencies than cocaine dependence, and for treatment effects for older adolescents.

The types of interventions included in this review were overwhelmingly manual-based interventions delivered by experienced therapists. Only one of the included studies studied 12-step programs in the self-help domain (Gossop et al., 2007), and this study was excluded from meta-analysis following risk of bias assessment. This review offers little evidence about the effectiveness of 12-step programs as delivered in self-help groups. There appears to be a discrepancy between how 12-step interventions are typically implemented in practice, in self-help groups, and the type of study designs and research evidence available for meta-analysis. All studies that were synthesised in meta-analysis were conducted in the US. In as far as population, setting, and delivery differs in other countries, the results should be interpreted in this light.

Economic considerations such as cost of treatment were not included in this review. None of the included studies supplied data on the cost of implementation. Because we cannot say that one treatment is better than another, based on the results of this review, considerations of cost may be particularly relevant.

6.2 IMPLICATIONS FOR RESEARCH

Although the main evidence presented in this review must be interpreted with caution, there are some important implications for research.

Compared to 12-step approaches, other psychosocial treatments appear to be neither worse nor better at reducing drug use. This conclusion has implications for research in this area. First, detecting small differences in the effectiveness of treatments will require large trials, or increased use of meta-analytical techniques to pool results across different trials. Second, a complementary approach is to more thoroughly establish a theory of change. This will be a first step allowing researchers to go beyond program names and towards understanding not only the active ingredients of treatments, but also understanding the importance of these ingredients for primary outcomes of interest. One approach to this challenge would be to develop intermediate treatment specific measures through which treatment is hypothesised to work. For example, 12-step programs are built up around a series of steps through which the patient gradually learns a set of

skills that allow them to deal with their addiction and achieve abstinence. Providing evidence on the gradual attainment of such skills and their influence on final or intermediate outcomes will be of great interest. Likewise, other psychosocial interventions specify competing mechanisms through which the patients will learn to master their addiction. Establishing whether such links exist will help us understand the commonalities between different programs and identify active ingredients. In addition, such a research program may involve complementary research methodologies of a more qualitative nature.

As our knowledge in this field gradually accumulates it will also be important for investigators to document and report potential prognostic factors e.g. social background, referral type (e.g. court-mandated, voluntary, etc.) and treatment motivation of participants. Meta-regression could then be applied to explore and generate hypotheses regarding positive and negative factors influencing treatment success. It will be of increasing importance that investigators report implementation costs in addition to the effects of intervention such that a cost-effectiveness analysis can be performed. This is particularly important in the context of drug abuse treatment because (a) treatment is typically very expensive, and (b) the negative impacts of drug abuse on friends, family and society may be great.

Many drug abuse treatments are delivered in a group therapy setting. It will be important for future research to adequately control for this design feature. Investigators should report or make available estimates of intra-cluster correlations and document group sizes. If such a practice was adopted, it would be possible to revisit reviews like this one and credibly correct the precision of effect size estimates for this design feature.

We need more evidence about the effects of 12-step programs for other populations. Increasing our knowledge of the effect of treatment for younger participants with drug dependence is an important area. Also increased knowledge of effects of 12-step programs for populations who are dependent on other drugs than cocaine is needed.

Finally, 12-step interventions to treat drug dependency are also offered in the form of self-help groups. It is plausible that this form of treatment attracts different populations than those studied in this review. For example participation in self-help groups, is often based more loosely on voluntary participation. Such 12-step services may be the next step after initial inpatient detoxification. This review found only one study which considered this setting. Future research may explore the effect of 12-step interventions in self-help groups. Identifying credible causal effects in these settings appears to be more challenging. Treatment intensity will be lower than in more institutionalised settings, and sample attrition may be a bigger threat to identification. Nevertheless, given the widespread practice and the lack of solid evidence on its effectiveness, this is an important area for future research.

7 References

7.1 REFERENCES TO INCLUDED STUDIES

Main trials marked with *

*Bissett, R. T. (2002). *Processes of change: Acceptance versus 12-step in polysubstance-abusing methadone clients.* ProQuest Information & Learning.

Hayes, S. C., Wilson, K. G., Gifford, E. V., Bissett, R., Piasecki, M., Batten, S. V., ... & Gregg, J. (2004). A preliminary trial of twelve-step facilitation and acceptance and commitment therapy with polysubstance-abusing methadone-maintained opiate addicts. *Behavior Therapy*, *35*(4), 667-688.

*Carroll, K. M., Nich, C., Ball, S. A., McCance, E., & Rounsavile, B. J. (1998). Treatment of cocaine and alcohol dependence with psychotherapy and disulfiram. *Addiction*, *93*(5), 713-727.

Carroll, K. M., Nich, C., Ball, S. A., McCance, E., Frankforter, T. L. & Rounsaville, B. J. (2000). One-year follow-up of disulfiram and psychotherapy for cocaine-alcohol users: sustained effects of treatment. *Addiction*, *95*(9), 1335-1349.

*Carroll, K. M., Nich, C., Shi, J. M., Eagan, D., & Ball, S. A. (2012). Efficacy of disulfiram and Twelve Step Facilitation in cocaine-dependent individuals maintained on methadone: a randomized placebo-controlled trial. *Drug & Alcohol Dependence*, *126*(1-2), 224-231.

*Gossop, M., Stewart, D., & Marsden, J., (2007). Attendance at Narcotics Anonymous and Alcoholics Anonymous meetings, frequency of attendance and substance use outcomes after residential treatment for drug dependence: a 5-year follow-up study. *Addiction*, *103*(1), 119-125.

*Higgins, S. T., Delaney, D. D., Budney, A. J., & Bickel, W. K. (1991). A behavioral approach to achieving initial cocaine abstinence. *The American Journal of Psychiatry*, *148*(9), 1218-1224.

*Maude-Griffin, P. M., Hohenstein, J. M., Humfleet, G. L., Reilly, P. M., Tusel, D. J., & Hall, S. M. (1998). Superior efficacy of cognitive-behavioral therapy for urban crack cocaine abusers: main and matching effects. *Journal of Consulting & Clinical Psychology*, *66*(5), 832-837.

*McKay, J. R., Alterman, A. I., Cacciola, J. S., Rutherford, M. J., O'Brien, C. P., & Koppenhaver, J. (1997). Group counseling versus individualized relapse prevention aftercare following intensive outpatient treatment for cocaine dependence: Initial results. *Journal of Consulting and Clinical Psychology, 65*(5), 778-788.

McKay, J. R., Alterman, A. I., Cacciola, J. S., O'Brien, C. P., Koppenhaver, J. M., & Shepard, D. S. (1999). Continuing care for cocaine dependence: Comprehensive 2-year outcomes. *Journal of Consulting and Clinical Psychology, 67*(3): 420-427.

*Petry, M, Weinstock, J., Alessi, S. M., Lewis, M. W., & Dieckhaus, K. (2010). Group-based randomized trial of contingencies for health and abstinence in HIV patients. *Journal of Consulting and Clinical Psychology, 78*(1), 89-97.

*Schottenfeld. R. S., Moore, B., & Pantalon, M. V. (2011). Contingency management with community reinforcement approach or twelve-step facilitation drug counseling for cocaine dependent pregnant women or women with young children. *Drug & Alcohol Dependence, 118*(1), 48-55.

*Wells, E. A., Peterson, P. L., Gainey, R. R., Hawkins, J. D., & Catalano, R. F. (1994). Outpatient treatment for cocaine abuse: a controlled comparison of relapse prevention and twelve-step approaches. *American Journal of Drug & Alcohol Abuse, 20*(1), 1-17.

7.2 REFERENCES TO EXCLUDED STUDIES

Aase, D. M., Jason, L. A., Ferrari, J. R., Li, Y., & Scott, G. (2014). Comorbid mental health and substance abuse issues among individuals in recovery homes: prospective environmental mediators. *Mental Health and Substance Use, 7*(2), 170-183.

Bergman, B. G., Hoeppner, B. B., Nelson, L. M., Slaymaker, V., & Kelly, J. F. (2015). The Effects of Continuing Care on Emerging Adult Outcomes Following Residential Addiction Treatment. *Drug and Alcohol Dependence,* 207-214.

Bogenschutz, M. P., Rice, S. L., Tonigan, J. S., Vogel, H. S., Nowinski, J., Hume, D., & Arenella, P. B. (2014). 12-step facilitation for the dually diagnosed: A randomized clinical trial. *Journal of substance abuse treatment, 46*(4), 403-411.

Bowen, S., Witkiewitz, K., Clifasefi, S. L., Grow, J., Chawla, N., Hsu, S. H., ... & Larimer, M. E. (2014). Relative efficacy of mindfulness-based relapse prevention, standard relapse prevention, and treatment as usual for substance use disorders: a randomized clinical trial. *JAMA psychiatry, 71*(5), 547-556.

Brooks, A. J., & Penn, P. E. (2003). Comparing treatments for dual diagnosis: Twelve-step and self-management and recovery training. *The American Journal of Drug and Alcohol Abuse, 29*(2), 359-383.

Brennan, P. I. (1998). Cognitive behavioral program vs. twelve-step program: Comparative effectiveness of two outpatient drug/alcohol treatment models. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 59*(6-B), 3049.

Brown, T. G., Seraganian, P., Tremblay, J., & Annis, H. (2002a). Process and outcome changes with relapse prevention versus 12-Step aftercare programs for substance abusers. *Addiction, 97*(6), 677-689.

Brown, T. G., Seraganian, P., Tremblay, J., & Annis, H. (2002b). Matching substance abuse aftercare treatments to client characteristics. *Addictive Behaviors, 27*(4), 585-604.

Chi, F. W., Weisner, C., Grella, C. E., Hser, Y. I., Moore, C., & Mertens, J. (2014). Does age at first treatment episode make a difference in outcomes over 11 years?. *Journal of substance abuse treatment, 46*(4), 482-490.

Donovan, D. M., Daley, D. C., Brigham, G. S., Hodgkins, C. C., Perl, H. I., Garrett, S. B., ... , & Kelly, T. M. (2013). Stimulant abuser groups to engage in 12-Step: A multisite trial in the National Institute on Drug Abuse Clinical Trials Network. *Journal of substance abuse treatment, 44*(1), 103-114.

Doyle, S. R., & Donovan, D. M. (2014). Applying an ensemble classification tree approach to the prediction of completion of a 12-step facilitation intervention with stimulant abusers. *Psychology of Addictive Behaviors, 28*(4), 1127-1143.

Fiorentine, R. (1999). After drug treatment: are 12-step programs effective in maintaining abstinence? *The American Journal of Drug and Alcohol Abuse, 25*(1), 93-116.

Henggeler, S. W., Pickrel, S. G., & Brondino, M, J. (1999). Multisystemic treatment of substance-abusing and-dependent delinquents: Outcomes, treatment fidelity, and transportability. *Mental Health Services Research, 1*(3), 171-184.

Henggeler, S. W., Clingempeel, W. G., Brondino, M. J., & Pickrel, S. G. (2002). Four-year follow-up of multisystemic therapy with substance-abusing and substance-dependent juvenile offenders. *Journal of the American Academy of Child & Adolescent Psychiatry, 41*(7), 868-874.

Lydecker, K. P., Tate, S. R., Cummins, K. M., McQuaid, J., Granholm, E., & Brown, S. A. (2010). Clinical outcomes of an integrated treatment for depression and substance use disorders. Psychology of Addictive Behaviors, 24(3), 453-465.

Majer, J. M., Jason, L. A., Aase, D. M., Droege, J. R., & Ferrari, J. R. (2013). Categorical 12-step involvement and continuous abstinence at 2 years. *Journal of substance abuse treatment, 44*(1), 46-51.

Manning, V., Best, D., Faulkner, N., Titherington, E., Morinan, A., Keaney, F., ... , & Strang, J. (2012). Does active referral by a doctor or 12-Step peer improve 12-Step meeting attendance? Results from a pilot randomised control trial. *Drug and alcohol dependence, 126*(1), 131-137.

McKay, J. R., Pettinati, H. M., Morrison, R., Feeley, M., Mulvaney, F. D., & Gallop, R. (2002). Relation of depression diagnoses to 2-year outcomes in cocaine-dependent patients in a randomized continuing care study. *Psychology of Addictive Behaviors, 16*(3), 225.

McKay, J. R., Lynch, K. G., Shepard, D. S., Ratichek, S., Morrison, R., Koppenhaver, J., & Pettinati, H. M. (2004). The effectiveness of telephone-based continuing care in the clinical management of alcohol and cocaine use disorders: 12-month outcomes. *Journal of Consulting and Clinical Psychology, 72*(6), 967-979.

McKay, J. R., Lynch, K. G., Shepard, D. S., & Pettinati, H. M. (2005). The effectiveness of telephone-based continuing care for alcohol and cocaine dependence: 24-month outcomes. *Archives of General Psychiatry, 62*(2), 199-207.

McKay, J. R., Lynch, K. G., Shepard, D. S., Morgenstern, J., Forman, R. F., & Pettinati, H. M. (2005). Do patient characteristics and initial progress in treatment moderate the effectiveness of telephone-based continuing care for substance use disorders? *Addiction, 100*(2), 216-226.

Morgan-Lopez, A. A., Saavedra, L. M., Hien, D. A., Campbell, A. N., Wu, E., & Ruglass, L. (2013). Synergy between seeking safety and twelve-step affiliation on substance use outcomes for women. *Journal of substance abuse treatment, 45*(2), 179-189.

Rosenblum, A., Matusow, H., Fong, C., Vogel, H., Uttaro, T., Moore, T. L., & Magura, S. (2014). Efficacy of dual focus mutual aid for persons with mental illness and substance misuse. *Drug and alcohol dependence, 135*, 78-87.

Wells, E. A., Donovan, D. M., Daley, D. C., Doyle, S. R., Brigham, G., Garrett, S. B., ... & Walker, R. (2014). Is level of exposure to a 12-step facilitation therapy associated with treatment outcome?. *Journal of substance abuse treatment, 47*(4), 265-274.

Worley, M. J., Tate, S. R., & Brown, S. A. (2012). Mediational relations between 12-Step attendance, depression and substance use in patients with comorbid substance dependence and major depression. *Addiction, 107*(11), 1974-1983.

Worley, M. J., Tate, S. R., McQuaid, J. R., Granholm, E. L., & Brown, S. A. (2013). 12-Step affiliation and attendance following treatment for comorbid substance dependence and depression: A latent growth curve mediation model. *Substance Abuse, 34*(1), 43-50.

7.3 ADDITIONAL REFERENCES

Alcoholics Anonymous (2012). Retrieved February 28, 2012, from http://www.aa.org.

Alcoholics Anonymous (2005). *Twelve steps and twelve traditions*. Alcoholics Anonymous. Electronic PDF version.

Brady, K. T., & Back, S. E. (2008). Women and addiction. In M. Galanter & H. D. Kleber (Eds.), *The American psychiatric publishing textbook of substance abuse treatment*. Washington, DC, American Psychiatric Publishing.

Cocaine Anonymous (2012). Retrieved February 28, 2012, from http://www.ca.org.

Cook, C. (1988). The Minnesota model in the management of drug and alcohol dependency: Miracle, method or myth? Part I. The philosophy and the programme. *British Journal of Addiction*, 83, 625-634.

Derogatis, L.R. (1983). SCL-90: Administration, Scoring and Procedures Manual-I for the Revised Version and other Instruments of the Psychopathology Rating Scale Series. Baltimore, MD: Johns Hopkins University School of Medicine, Clinical Psychometrics Research Unit.

European Monitoring Centre for Drugs and Drug Addiction, EMCDDA (2010). *Annual report 2009: The state of the drugs problem in Europe*. Luxembourg: Publications office of the European Union.

Farabee, D., Prendergast, M., & Anglin, M. D. (1998). The Effectiveness of Coerced Treatment for Drug-Abusing Offenders. *Federal Probation*, 62, 3.

Ferri, M., Amato, L., & Davoli, M. (2006). Alcoholics Anonymous and other 12-step programmes for alcohol dependence. *The Cochrane Library*, 3.

Finney, J. W., Noyes, C. A., Coutts, A. I., & Moos, R. H. (1998). Evaluating substance abuse treatment process models: I. Changes on proximal outcome variables during 12-Step and cognitive-behavioral treatment. *Journal of Studies on Alcohol, 59*, 371-380.

Fiorentine, R. (1999). After drug treatment: Are 12-step programs effective in maintaining abstinence. *American Journal of Drug and Alcohol Abuse, 25(1)*, 93-116.

Forman, R. F., Bovasso, G., & Woody, G. (2001). Staff beliefs about addiction treatment. *Journal of Substance Abuse Treatment, 21*, 1-19.

Galanter, M., & Kleber, H. D. (Eds.) (2008). *The American psychiatric publishing textbook of substance abuse treatment*. Washington, DC: American Psychiatric Publishing.

Greenfield, S. F., & Hennessy, G. (2008). Assessment of the patient. In M. Galanter & H. D. Kleber (Eds.), *The American psychiatric publishing textbook of substance abuse treatment*. Washington, DC: American Psychiatric Publishing.

Griffith, J.D., Rowan-Szal, G.A., Roark, RR., & Simpson, D.D. (2000). Contingency management in outpatient methadone treatment: a meta-analysis. *Drug and Alcohol Dependence, 58*, 55-66.

Higgins, J. P. T., & Green, S. (Eds.) (2008). *Cochrane handbook for systematic reviews of interventions version 5.0.2* [updated September 2009]. The Cochrane Collaboration (available from http://www.cochrane-handbook.org).

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21*, 1539-1558.

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ: British Medical Journal, 327*, 557-560.Hornbacher, M. (2010). Sane: mental illness, addiction, and the twelve steps. Center City, MN: Hazelden.

Humphreys, K., & Noke, J. M. (1997). The influence of posttreatment mutual help group participation on the friendship networks of substance abuse patients. *American Journal of Community Psychology, 25*, 1-16.

Jennings, J., & Alcoholics Anonymous (1990). *The Little Red Book*. Center City, MN: Hazelden.

Kelly, J. F., Finney, J. W., & Moos, R. (2005). Substance use disorder patients who are mandated to treatment: Characteristics, treatment process, and 1-and 5-year outcomes. *Journal of Substance Abuse Treatment*, 28(3), 213-223.

Kessler, R. C., Crum, R. M., Warner, L. A., Nelson, C. B., Shulenberg, J., & Anthony, J. C. (1997). Lifetime co-occurrence of DSM-III-R alcohol abuse and dependence with other psychiatric disorders in the National Comorbidity Survey. *Archives of General Psychiatry, 54*, 313-321.

Kiluk, B. D., Nich, C., Witkiewitz, K., Babuscio, T. A., & Carroll, K. M. (2014). What happens in treatment doesn't stay in treatment: Cocaine abstinence during treatment is associated with fewer problems at follow-Up. *Journal of Consulting and Clinical Psychology*, *82*(4), 619-627.

Kish, L. (1965). *Survey Sampling*. New York: Wiley

Kline, A. (1997). Profiles of criminal-justice clients in drug treatment: Implications for intervention. *Addictive behaviors*, 22(2), 263-268.

Labouvie, E. (1996). Maturing out of substance use: Selection and self correction. *Journal of Drug Issues, 26(2)*, 457-476.

Laudet, A. B. (2008). The impact of Alcoholics Anonymous on other substance abuse related Twelve Step programs. *Recent Developments in Alcoholism, 18*, 71-89.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Applied Social Research Methods Series, v. 49.

Marsch, L.A. (1998). The efficacy of methadone maintenance inttervenions in reducing illicit opiate use, HIV risk behaviors and criminality: a meta-analysis. Addiction, 93, 515-532.

Mäkelä, K., Arminen, I., Bloomfield, K., Eisenbach-Stangl, I., Hermersson Bergmark, K., Kurube, N., Mariolini, N., ... Zielinksi, A. (1996). *Alcoholics Anonymous as a mutual-help movement: A study in eight societies*. Wisconsin University Press, Madison, Wisconsin, USA.

Marijuana Anonymous (2012). Retrieved February 28, 2012, from http://www.marijuana-anonymous.org.

MATCH (1997). Matching alcoholism to client heterogeneity: Project MATCH posttreatment drinking outcomes. *Journal of Studies on Alcohol and Drugs, 58(1),* 7-29.

McLellan, A. T., Luborsky, L., Woody, G. E., & O'Brien, C. P. (1980). An improved diagnostic evaluation instrument for substance abuse patients. The addiction severity index. *Journal of Nervous and Mental Disorders, 168(1),* 26-33.

Mercer, D. E., & Woody, G. E. (1999). Contributions of the 12-Step Approach. In *An individual drug counselling approach to treat cocaine addiction*. National Institute on Drug Abuse, Division of Clinical and Services Research.

Moos, R. H., & Timko, C. (2008). *Outcome research on 12-Step and other self-help programs.* In M. Galanter & H. D. Kleber (Eds.), The American psychiatric publishing textbook of substance abuse treatment. Washington, DC:American Psychiatric Publishing.

Morgenstern, J., Bux, D., & Labouvie, E. (2002). Examining mechanisms of action in 12-step treatment: The role of 12-step cognitions. *Journal on Studies of Alcohol and Drugs, 63,* 665-671.

Narcotics Anonymous (2008). *Narcotics Anonymous –Basic text*. Sixth edition. Narcotics Anonymous World Services, Inc.

Narcotics Anonymous (2010). Information about NA. Retrieved February 28, 2012), from http://www.na.org/admin/include/spaw2/uploads/pdf/PR/Information_about_NA.pdf.

Narcotics Anonymous (2012). Retrieved February 28, 2012, from http://www.na.org.

Office of National Drug Control Policy, ONDCP (2000). *Drug-related crime*. Office of National Drug Control Policy, Drug Policy Information Clearinghouse, fact sheet.

Pals, S. L., Murray, D. M., Alfano, C. M., Shadish, W. R., Hannan, P. J., & Baker, W. L. (2008). Individually randomized group treatment trials: a critical appraisal of frequently used design and analytic approaches. *American Journal of Public Health, 98*(8), 1418-1424.

Perron, B. E., & Bright, C. L. (2008). The influence of legal coercion on dropout from substance abuse treatment: Results from a national survey. *Drug and alcohol dependence,* 92(1), 123-131.

Pigott, T. D. (2009). Handling missing data. In H. Cooper & L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 399-416). New York: Russell Sage Foundation.

Pills Anonymous (2012). Retrieved February 28, 2012, from http://www.pillsanomynous.org.

Prendergast, M.L., Podus, D., Chang, E, & Urada, D. (2002). The effectiveness of drug abuse treatment: a meta-analysis of comparison group studies. *Drug and Alcohol Dependence, 67,* 53-72.

Prochaska, J. O., & DiClemente, C. C. (1982). Transtheoretical therapy: Toward a more integrative model of change. *Psychotherapy: Theory, Research & Practice*, 19(3), 276.

Reeves, B.C., Deeks J.J., Higgins J.P.T., & Wells G.A. (2011). *Chapter 13: Including non-randomized studies*. In: Higgins J.P.T., & Green S. (Eds.), Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from www.handbook.cochrane.org.

Review Manager Version 5.3. [Computer software]. (2014). Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration.

Ries, R. K., Galanter, M., & Tonigan, J. S. (2008). *Twelve-step facilitation*. In M. Galanter & H. D. Kleber (Eds.), The American psychiatric publishing textbook of substance abuse treatment. Washington, DC: American Psychiatric Publishing.

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychol. Bulletin, 86*, 638-641.

Ross, S. (2008). *The mentally ill substance abuser*. In M. Galanter & H. D. Kleber (Eds.), The American psychiatric publishing textbook of substance abuse treatment. Washington, DC: American Psychiatric Publishing.

Sánchez-Meca, J., Márin-Martínes, F., & Chacón-Moscoso, S. (2003). Effect-size indices for dichotomized outcomes in meta-analysis. *Psychological Methods, 8(4)*, 448-467.

Shadish, W. R., & Cook, T. D. (2009). The Renaissance of Field Experimentation in Evaluating Interventions. *Annual Review of Psychology, 60(1)*, 607-629.

Shelton, K. H., Taylor, P. J., Bonner, A., & van den Bree, M. (2009). Risk factors for homelessness: Evidence from a population-based study. *Psychiatric Services, 60*, 465-472.

Silbert, M. H., Pines, A. M., & Lynch, T. (1982). Substance abuse and prostitution. *Journal of Psychoactive Drugs, 14(3)*, 193-7.

Sobell, L. C., Sobell, M. B., Buchan, G., Cleland, P. A., Fedoroff, I., & Leo, G. I. (1996). The reliability of the Timeline Followback method applied to drug, cigarette, and cannabis use. Paper presented at the 30th Annual Meeting of the Association for Advancement of Behavior Therapy, New York, NY, November 1996.

Spooner, C., & Hetherington, K. (2004). *Social determinants of drug use*. National Drug and Alcohol Research Centre, University of New South Wales, Technical Report Number 228.

StataCorp. (2013). Stata Statistical Software: Release 13 [Computer software]. College Station, TX: StataCorp LP.

Stinson, F. S., Grant, B. F., Dawson, D. A., Ruan, W. J., Huang, B., & Saha, T. (2005). Comorbidity between DSM-IV alcohol and specific drug use disorders in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Drug and alcohol dependence*, 80(1), 105-116.

Straussner, S., & Spiegel, B. (1996). An analysis of 12-step programs for substance abusers from a developmental perspective. *Clinical Social Work Journal, 24(3),* 299-309.

Tonigan, J. S., & Connors, G. J. (2008). *Psychological mechanisms in Alcoholics Anonymous*. In M. Galanter & H. D. Kleber (Eds.), The American psychiatric publishing textbook of substance abuse treatment. Washington, DC: American Psychiatric Publishing.

Tyrer, P., Nur, U., Crawford, M., Karlsen, S., McLean, C., Rao, B. & Johnson, T. (2005). The Social Functioning Questionnaire: a rapid and robust measure of perceived functioning. International *Journal of Social Psychiatry, 51*(3), 265-75.

United Nations Office on Drugs and Crime, UNODC (2010), *World Drug Report 2010*. United Nations Publication.

Witkiewitz, K., & Marlatt, A. G. (2011). Behavioral therapy across the spectrum. *Alcohol Research & Health, 33*, 4.

8 About this review

8.1 ACKNOWLEDGEMENTS

The review authors would like to thank the Campbell methods peer referees, Dr Nick Huband, Professor Cathy Bennett, Professor Brandy R. Maynard, Dr Jane Dennis, Ms Catriona Shatford, Dr Emily Tanner-Smith, and the anonymous external content and methods peer referees, for valuable and insightful comments on methods and content during the stage of writing the protocol and the final review report. We are grateful to the study authors of Carroll et al. (1998, 2000, 2012) for their kind response and help in providing additional study data, particularly Professor Kathleen Carroll, and Dr Theresa Babuscio.

Thanks to Head of SFI Campbell, PhD Mette Deding, for continued support and efforts to realise this review. Majken Mosegaard Svendsen and Anne Sofie Due Knudsen helped draft the protocol, and Majken assisted with descriptive and numerical coding. Ida Scheel Rasmussen provided excellent assistance at key stages in the production of this review.

8.2 DIFFERENCES BETWEEN PROTOCOL AND REVIEW

In the protocol we stated that we would synthesise outcomes by calendar time as follows:

- Short-term effects (less than 6 months after enrolment into treatment).
- Medium-term effects (6 months to less than 12 months after enrolment into treatment).
- Long-term effects (12 months or more after enrolment into treatment).

This decision was made because we expect that the majority of included studies would be based on treatments in self-help groups. Only one included study had treatments of this nature (Gossop et al., 2007). The remaining studies were provided in outpatient settings by therapists, typically manual-based interventions such as TSF. Since these interventions had a fixed duration, which varied between studies, and studies frequently reported on the primary outcome during, post, and at follow-up, we decided to follow the included literature in synthesising the available evidence, during, post-treatment, and at follow-up.

We had planned to do moderator analyses of effect sizes. Because of the small number of studies uncovered during the search process, we did not perform meta-regression.

In the protocol we had planned to do regression based tests for publication bias. However we did not find a sufficient number of studies to perform these tests.

8.3 REVIEW AUTHORS

Lead review author:

The lead author is the person who develops and co-ordinates the review team, discusses and assigns roles for individual members of the review team, liaises with the editorial base and takes responsibility for the on-going updates of the review.

Name:	**Martin Bøg**
Title:	PhD
Affiliation:	SFI Campbell/Trials Unit
Address:	Herluf Trollesgade 11
Postal Code:	DK-1052 Copenhagen K
Country:	Denmark
Phone:	+45 3369 7741
Email:	martin.bog@gmail.com

Co-author(s):

Name:	**Trine Filges**
Title:	PhD
Affiliation:	SFI Campbell/Trials Unit
Address:	Herluf Trollesgade 11
Postal Code:	DK-1052 Copenhagen K
Country:	Denmark
Email:	tif@sfi.dk

Co-author(s):

Name:	**Lars Brännström**
Title:	PhD
Affiliation:	Department of Social Work, Stockholm University
Address:	Stockholms universitet Institutionen för socialt arbete - socialhögskolan
City, State, Province or County:	
Postal Code:	SE-106 91 Stockholm
Country:	Sweden
Phone:	+46 (0)8-16 48 59
Email:	lars.brannstrom@socarb.su.se

Co-author(s):

Name:	**Anne-Marie Klint Jørgensen**
Title:	Information specialist
Affiliation:	SFI Campbell/Trials Unit
Address:	Herluf Trollesgade 11
Postal Code:	DK-1052 Copenhagen K
Country:	Denmark
Email:	amk@sfi.dk

Co-author(s):

Name:	**Maja Kärrman Fredriksson**
Title:	Information specialist
Affiliation:	Socialstyrelsen Avdelningen för kunskapsstyrning Kunskapsöversikter
Postal Code:	SE-106 30 Stockholm
Country:	Sweden
Phone:	+46 752473159
Email:	maja.karrman-fredriksson@socialstyrelsen.se

8.4 ROLES AND RESPONSIBILITIES

- Content: Martin Bøg, Lars Brännström
- Systematic review methods: Trine Filges, Martin Bøg
- Statistical analysis: Trine Filges, Martin Bøg
- Information retrieval: Anne-Marie Klint Jørgensen, Maja Kärrman Fredriksson, Bjørn Christian Viinholt Nielsen
- Coding:
  - Descriptive: Martin Bøg, Ida Scheel Rasmussen, Majken Mosegaard Svendsen
  - Numerical: Martin Bøg, Trine Filges, Ida Scheel Rasmussen, Majken Mosegaard Svendsen
  - Risk of Bias: Martin Bøg, Trine Filges, Ida Scheel Rasmussen
- Screening: Anne-Sofie Due Knudsen, Simon Helth Filges, Pia Vang Hansen, Stine Lian Olsen, Madina Saidj, Malan Ó á Dunga, Marcel Mirzaei-Fard, Katrine Strøjer Madsen, Freja Jørgensen, Sjúrdur Zachariasson, Trine Filges, Martin Bøg, Rasmus Henriksen Klokker

8.5 SOURCES OF SUPPORT

SFI Campbell

8.6 DECLARATIONS OF INTEREST

The authors have no vested interest in the outcomes of this review, nor any incentive to represent findings in a biased manner.

8.7 AUTHOR DECLARATION

Authors' responsibilities

By completing this form, you accept responsibility for maintaining the review in light of new evidence, comments and criticisms, and other developments, and updating the review at least once every five years, or, if requested, transferring responsibility for maintaining the review to others as agreed with the Coordinating Group. If an update is not submitted according to agreed plans, or if we are unable to contact you for an extended period, the relevant Coordinating Group has the right to propose the update to alternative authors.

Publication in the Campbell Library

The Campbell Collaboration places no restrictions on publication of the findings of a Campbell systematic review in a more abbreviated form as a journal article either before or after the publication of the monograph version in *Campbell Systematic Reviews*. Some journals, however, have restrictions that preclude publication of findings that have been, or will be, reported elsewhere, and authors considering publication in such a journal should be aware of possible conflict with publication of the monograph version in *Campbell Systematic Reviews*. Publication in a journal after publication or in press status in *Campbell Systematic Reviews* should acknowledge the Campbell version and include a citation to it. Note that systematic reviews published in *Campbell Systematic Reviews* and co-registered with the Cochrane Collaboration may have additional requirements or restrictions for co-publication. Review authors accept responsibility for meeting any co-publication requirements.

9 Tables

9.1 SEARCH TERM BY DATABASE

Bibliotek.dk

Searched January 2011.

Update July 2013, August 2015, and final update September 2016.

Search	Term	Totals
s1	(narcotics anonymous) eller (cocaine anonymous?) eller (crystal meth anonymous?) eller (pills anonymous?) eller (marijuana anonymous?) eller (heroin anonymous?) eller (12 step) eller (selvhjælpsgrupper) eller (selfhelp group?) eller (minnesota model?) eller (minnesota program?) eller (minnesota behandling?)	51
s2	(narcotics anonymous) eller (cocaine anonymous?) eller (crystal meth anonymous?) eller (pills anonymous?) eller (marijuana anonymous?) eller (heroin anonymous?) eller (12 step) eller (selvhjælpsgrupper) eller (selfhelp group?) eller (minnesota model?) eller (minnesota program?) eller (minnesota behandling?) – Limited to 2015-current	2

Bibsys (Bibliotekbasen, ForskDok publikasjoner, ForskDok prosjekter)

Searched January 2011.

Updates in September 2011, July 2013, August 2015. In the newest update of the review (September 2016), the access to BIBSY (now Oria) was limited. The documented search strategy reflects searches until August 2015.

Search number	Term	Totals
s1	Anonyme narkomane eller narcotics anonymous	13

s2	Cocaine Anonymous	4
s3	marijuana Anonymous eller anonyme og hasj og avhengig?	2
s4	Heroin Anonymous	1
s5	Anonyme narkomane eller narcotics anonymous eller Cocaine Anonymous eller marijuana Anonymous eller anonyme og hasj og avhengig? eller Heroin Anonymous	15
s6	Selvhjelpsgruppe	7
s7	støttegruppe	12
s8	Twelve-step? eller tolv trinn? eller tolv tradisjoner	31
s9	12-step? eller 12 trinn?	44
s10	emne = selvhjelpsgruppe?	160
s11	model? eller program? eller behandling? eller rehabilit? eller kur eller terapi eller detox eller avrus eller recovery eller intervensjon? Eller metode? og minnesota	140
s12	TSF	15
s13	Selvhjelpsgruppe eller støttegruppe eller Twelve-step? eller tolv trinn? eller tolv tradisjoner eller 12-step? eller 12 trinn? eller emne =selvhjelpsgruppe? eller model? eller program? eller behandling? Eller rehabilit? eller kur eller terapi eller detox eller avrus eller recovery eller intervensjon? eller metode? og minnesota eller TSF	398
s14	emne = Amphetamine Related Disorders	48
s15	emne = cocaine related disorders	181
s16	emne = marijuana abuse	107
s17	emne = opioid related disorders	713
s18	emne = substance abuse intravenous	94
s19	emne = Substance Withdrawal Syndrome eller abstinenser	31
s20	emne = heroin abuse	55
s21	emne = Amphetamine Related Disorders eller emne = cocaine related disorders eller emne = marijuana abuse eller emne = opioid related disorders eller emne = substance abuse intravenous eller emne = Substance Withdrawal Syndrome eller abstinenser eller emne = heroin	1025
s22	emne = opium	116
s23	emne = heroin	277
s24	emne = cannabinoid? eller emne = cannabis	372
s25	emne = Marijuana Smoking	88
s26	emne = cocaine	322
s27	emne = methamphetamine	27
s28	emne = Amphetamine	73

s29	emne = Designer Drugs	71
s30	emne = opium eller emne = heroin eller emne = cannabinoid? eller emne = cannabis eller emne = Marijuana Smoking eller emne = cocaine eller emne = methamphetamine eller emne = Amphetamine eller emne = Designer Drugs	1059
s31	stimulan?	238
s32	Opium	310
s33	Heroin	388
s34	Crack	607
s35	Cocaine eller kokain	536
s36	Methamphetamine eller metamfetamin? eller amphetamin? eller amfetamin?	267
s37	Ecstasy	314
s38	stimulan? eller Opium eller Heroin eller Crack eller Cocaine eller kokain eller Methamphetamine eller metamfetamin? eller amphetamin? Eller amfetamin? eller Ecstasy	2272
s39	Cannabis eller marijuana eller marihuana eller hash eller hasj eller hashish	842
s40	drug? eller stof? og designer	72
s41	Narko?	4363
s42	Drug? eller stof?	22604
s43	stof? eller drug? eller stimulan? eller narko? og misbruk? eller brug? Eller avheng?	697
s44	Anonyme narkomane eller narcotics anonymous eller Cocaine Anonymous eller marijuana Anonymous eller anonyme og hasj og avhengig? eller Heroin Anonymous eller Selvhjelpsgruppe eller støttegruppe eller Twelve-step? eller tolv trinn? eller tolv tradisjoner eller 12-step? eller 12 trinn? eller emne = selvhjelpsgruppe? eller model? Eller program? eller behandling? eller rehabilit? eller kur eller terapi eller detox eller avrus eller recovery eller intervensjon? eller metode? Og minnesota eller TSF	411
s45	emne = Amphetamine Related Disorders eller emne = cocaine related disorders eller emne = marijuana abuse eller emne = opioid related disorders eller emne = substance abuse intravenous eller emne = Substance Withdrawal Syndrome eller abstinenser eller emne = heroin abuse og Anonyme narkomane eller narcotics anonymous eller Cocaine Anonymous eller marijuana Anonymous eller anonyme og hasj og avhengig? eller Heroin Anonymous eller Selvhjelpsgruppe eller	7

	støttegruppe eller Twelve-step? eller tolv trinn? eller tolv tradisjoner eller 12-step? eller 12 trinn? eller emne = selvhjelpsgruppe? eller model? Eller program? eller behandling? eller rehabilit? eller kur eller terapi eller detox eller avrus eller recovery eller intervensjon? eller metode? Og minnesota eller TSF	
s46	emne = opium eller emne = heroin eller emne = cannabinoid? eller emne = cannabis eller emne = Marijuana Smoking eller emne = cocaine eller emne = methamphetamine eller emne = Amphetamine eller emne = Designer Drugs og Anonyme narkomane eller narcotics anonymous eller Cocaine Anonymous eller marijuana Anonymous eller anonyme og hasj og avhengig? eller Heroin Anonymous eller Selvhjelpsgruppe eller støttegruppe eller Twelve-step? eller tolv trinn? eller tolv tradisjoner eller 12-step? eller 12 trinn? eller emne = selvhjelpsgruppe? eller model? Eller program? eller behandling? eller rehabilit? eller kur eller terapi eller detox eller avrus eller recovery eller intervensjon? eller metode? Og minnesota eller TSF	3
s47	stimulan? eller Opium eller Heroin eller Crack eller Cocaine eller kokain eller Methamphetamine eller metamfetamin? eller amphetamin? Eller amfetamin? eller Ecstasy og Anonyme narkomane eller narcotics anonymous eller Cocaine Anonymous eller marijuana Anonymous eller anonyme og hasj og avhengig? eller Heroin Anonymous eller Selvhjelpsgruppe eller støttegruppe eller Twelve-step? eller tolv trinn? eller tolv tradisjoner eller 12-step? eller 12 trinn? eller emne = selvhjelpsgruppe? eller model? eller program? eller behandling? Eller rehabilit? eller kur eller terapi eller detox eller avrus eller recovery eller intervensjon? eller metode? og minnesota eller TSF	6
s48	Cannabis eller marijuana eller marihuana eller hash eller hasj eller hashish og Anonyme narkomane eller narcotics anonymous eller Cocaine Anonymous eller marijuana Anonymous eller anonyme og hasj og avhengig? eller Heroin Anonymous eller Selvhjelpsgruppe eller støttegruppe eller Twelve-step? eller tolv trinn? eller tolv tradisjoner eller 12-step? eller 12 trinn? eller emne = selvhjelpsgruppe? eller model? Eller program? eller behandling? eller rehabilit? eller kur eller terapi eller detox eller avrus eller recovery eller intervensjon? eller metode? Og minnesota eller TSF	5
s49	drug? eller stof? og designer og Anonyme narkomane eller narcotics anonymous eller Cocaine Anonymous eller marijuana Anonymous eller anonyme og hasj og avhengig? eller Heroin Anonymous eller	0

	Selvhjelpsgruppe eller støttegruppe eller Twelve-step? eller tolv trinn? eller tolv tradisjoner eller 12-step? eller 12 trinn? eller emne = selvhjelpsgruppe? eller model? eller program? eller behandling? eller rehabilit? eller kur eller terapi eller detox eller avrus eller recovery eller intervensjon? eller metode? og minnesota eller TSF	
s50	Narko? og Anonyme narkomane eller narcotics anonymous eller Cocaine Anonymous eller marijuana Anonymous eller anonyme og hasj og avhengig? eller Heroin Anonymous eller Selvhjelpsgruppe eller støttegruppe eller Twelve-step? eller tolv trinn? eller tolv tradisjoner eller 12-step? eller 12 trinn? eller emne = selvhjelpsgruppe? eller model? Eller program? eller behandling? eller rehabilit? eller kur eller terapi eller detox eller avrus eller recovery eller intervensjon? eller metode? Og minnesota eller TSF	14
s51	Drug? eller stof? og Anonyme narkomane eller narcotics anonymous eller Cocaine Anonymous eller marijuana Anonymous eller anonyme og hasj og avhengig? eller Heroin Anonymous eller Selvhjelpsgruppe eller støttegruppe eller Twelve-step? eller tolv trinn? eller tolv tradisjoner eller 12-step? eller 12 trinn? eller emne = selvhjelpsgruppe? eller model? Eller program? eller behandling? eller rehabilit? eller kur eller terapi eller detox eller avrus eller recovery eller intervensjon? eller metode? Og minnesota eller TSF	28
s52	stof? eller drug? eller stimulan? eller narko? og misbruk? eller brug? Eller avheng? og Anonyme narkomane eller narcotics anonymous eller Cocaine Anonymous eller marijuana Anonymous eller anonyme og hasj og avhengig? eller Heroin Anonymous eller Selvhjelpsgruppe eller støttegruppe eller Twelve-step? eller tolv trinn? eller tolv tradisjoner eller 12-step? eller 12 trinn? eller emne = selvhjelpsgruppe? eller model? Eller program? eller behandling? eller rehabilit? eller kur eller terapi eller detox eller avrus eller recovery eller intervensjon? eller metode? Og minnesota eller TSF	5
S53	stof? eller drug? eller stimulan? eller narko? og misbruk? eller brug? Eller avheng? og Anonyme narkomane eller narcotics anonymous eller Cocaine Anonymous eller marijuana Anonymous eller anonyme og hasj og avhengig? eller Heroin Anonymous eller Selvhjelpsgruppe eller støttegruppe eller Twelve-step? eller tolv trinn? eller tolv tradisjoner eller 12-step? eller 12 trinn? eller emne = selvhjelpsgruppe? eller model? Eller program? eller behandling? eller rehabilit? eller kur eller terapi eller detox eller avrus eller	0

recovery eller intervensjon? eller metode? Og minnesota eller TSF – Limited to 2015-current.

CINAHL

Searched January 2010

Update September 2011, July 2013, August 2015, no access in final update September 2016.

Search number	Term	Totals
s1	TI (narcotics anonymous or Cocaine Anonymous or Crystal Meth Anonymous or Pills Anonymous or Marijuana Anonymous or Heroin Anonymous) or AB (narcotics anonymous or Cocaine Anonymous or Crystal Meth Anonymous or Pills Anonymous or Marijuana Anonymous or Heroin Anonymous) Search modes	29
s2	TI Self-Help N1 group* or AB Self-Help N1 group*	491
s3	TI Support* n1 group* or AB Support* n1 group*	3605
s4	TI (12-step* or 12 n1 step*) or AB (12-step* or 12 n1 step*)	259
s5	TI (twelve-step* or twelve n1 step*) or AB (twelve-step* or twelve n1 step*)	89
s6	(MM "Support Groups")	3053
s7	TI (Recover* n1 group*) and AB (Recover* n1 group*)	5
s8	TI Minnesota n3 model* or AB Minnesota n3 model*	18
s9	TI Minnesota n3 program* or AB Minnesota n3 program* Search Screen - Advanced Search	77
s10	TI Minnesota n3 treatment* or AB Minnesota n3 treatment* Search modes Boolean/Phrase Interface - EBSCOhost Search Screen - Advanced Search	20
s11	TI Minnesota n3 rehab* or AB Minnesota n3 rehab* Search modes - Boolean/Phrase Interface - EBSCOhost Search Screen - Advanced Search	7
s12	TI Minnesota n3 cure* or AB Minnesota n3 cure* Search modes -	0

	Search Screen - Advanced Search	
s13	TI Minnesota n3 therap* or AB Minnesota n3 therap*	8
s14	TI Minnesota n3 detox* or AB Minnesota n3 detox*	0
s15	TI Minnesota n3 recover* or AB Minnesota n3 recover*	0
s16	TI Minnesota n3 intervent* or AB Minnesota n3 intervent*	14
s17	TI Minnesota n3 method* or AB Minnesota n3 method*	46
s18	S8 or S9 or S10 or S11 or S12 or S13 or S14 or S15 or S16 or S17	171
s19	S1 or S2 or S3 or S4 or S5 or S6 or S7 or S18	1641
s20	TI Drug* or AB Drug*	80573
s21	TI Substance* or AB Substance*	15339
s22	TI Stimulan* or AB Stimulan*	1186
s23	TI Narcotic* or AB Narcotic*	1160
s24	TI (Opium or Heroin or Crack or Cocaine* or Methamphetamine* or Amphetamin*or Ecstasy or Fantasy or Cannabis or Marijuana or Hashish) or AB (Opium or Heroin or Crack or Cocaine* or Methamphetamine* or Amphetamin*or Ecstasy or Fantasy or Cannabis or Marijuana or Hashish)	7495
s25	TI Designer n1 drug* or AB Designer n1 drug*	54
s26	S20 or S21 or S22 or S23 or S24 or S25	96576
s27	TI substance* n3 abus* or AB substance* n3 abus*	6328
s28	TI substance* n3 us* or AB substance* n3 us*	5696
s29	TI substance* n3 addict* or AB substance* n3 addict	253

s30	TI substance* n3 depend* or AB substance* n3 depend*	701
s31	TI substance* n3 misus* or AB substance* n3 misus*	669
s32	S27 or S28 or S29 or S30 or S31	11454
s33	TI drug* n3 abus* or AB drug* n3 abus*	3252
s34	TI drug* n3 misus* or AB drug* n3 misus*	502
s35	TI drug* n3 addict* or AB drug* n3 addict*	919
s36	TI drug* n3 depend* or AB drug* n3 depend*	1151
s37	TI drug* n3 us* or AB drug* n3 us*	17322
s38	S33 or S34 or S35 or S36 or S37	21091
s39	TI stimulan* n3 abus* or AB stimulan* n3 abus*	58
s40	TI stimulan* n3 us* or AB stimulan* n3 us*	322
s41	TI stimulan* n3 misus* or AB stimulan* n3 misus*	11
s42	TI stimulan* n3 addict* or AB stimulan* n3 addict* Search modes -	16
s43	TI stimulan* n3 depend* or AB stimulan* n3 depend*	28
s44	S39 or S40 or S41 or S42 or S43	394
s45	TI Narcotic* n3 abus* or AB Narcotic* n3 abus*	22
s46	TI Narcotic* n3 us* or AB Narcotic* n3 us*	266
s47	TI Narcotic* n3 misus* or AB Narcotic* n3 misus*	2

s48	TI Narcotic* n3 addict* or AB Narcotic* n3 addict*	35
s49	TI Narcotic* n3 depend* or AB Narcotic* n3 depend*	15
s50	S45 or S46 or S47 or S48 or S49 Boolean/Phrase Interface - EBSCOhost	331
s51	S32 or S38 or S44 or S50 Boolean/Phrase Interface - EBSCOhost Search Screen - Advanced Search	30289
s52	TI TSF or AB tsf	66
s53	(S19 or S52)	6189
s54	S26 and S53	429
s55	S51 and S53	340
s56	(S54 or S55)	429
s57	(S54 or S55) Limiters - Published Date from: 20100901-20110931 Search Screen - Advanced Search	21

Cochrane

Search February 2011. Update September 2011, July 2013, August 2015, and final update September 2016.

Search number	Terms	Totals
s1	(narcotics anonymous):ti,ab,kw	9
s2	(Cocaine Anonymous):ti,ab,kw	7
s3	(Crystal Meth Anonymous):ti,ab,kw	0
s4	(Pills Anonymous):ti,ab,kw	1
s5	(Marijuana Anonymous):ti,ab,kw	2

S6	(Heroin Anonymous):ti,ab,kw	1
s7	(#1 OR #2 OR #3 OR #4 OR #5 OR #6)	16
s8	(Self-Help adj1 group*):ti,ab,kw	0
s9	(Support* adj1 group*).:ti,ab,kw	0
s10	(twelve-step* or twelve near/1 step*):ti,ab,kw	45
s11	(12-step* or 12 near/1 step*):ti,ab,kw	67
s12	Self-Help Groups/	1420
s13	(Recover* near/1 group*):ti,ab,kw	204
s14	(Minnesota near/3 (model* or program* or treatment*or Rehab* or cure* or therap* or detox* or recover* or intervent* or method*)):ti,ab,kw	149
s15	(TSF):ti,ab,kw	42
s16	(#8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15)	1858
s17	Amphetamine-Related Disorders/	100
s18	Cocaine-Related Disorders/	473
s19	Marijuana Abuse/	311
s20	Opioid-Related Disorders/	584
s21	Substance Withdrawal Syndrome/	1838
s22	Heroin Dependence/	709
s23	(#17 OR #18 OR #19 OR #20 OR #21 OR #22)	3428
s24	Opium/	156
s25	Heroin/	1057
s26	exp Cannabinoids/ or Cannabis/	686
s27	Marijuana Smoking/	317
s28	exp cocaine/	55
s29	Methamphetamine/	281
s30	Amphetamine/	1027
s31	Designer Drugs/	29
s32	(Drug*):ti,ab,kw	276295
s33	(Substance*):ti,ab,kw	9439
s34	(Stimulan*):ti,ab,kw	2477
s35	(Narcotic*):ti,ab,kw	3812
s36	(Opium.):ti,ab,kw	127
s37	(Heroin):ti,ab,kw	989
s38	(Crack):ti,ab,kw	225
s39	(Cocaine*):ti,ab,kw	1749
s40	(Methamphetamine* or Amphetamin*).:ti,ab,kw	0

s41	(Ecstasy):ti,ab,kw	76
s42	(Fantasy):ti,ab,kw	104
s43	(Cannabis or Marijuana or Hashish).:ti,ab,kw	0
s44	(Designerdrug* or (designer near/1 drug*)):ti,ab,kw	7
s45	(#24 OR #25 OR #26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32 OR #33 OR #34 OR #35 OR #36 OR #37 OR #38 OR #39 OR #40 OR #41 OR #42 OR #43 OR #44)	281719
s46	(substance* or drug* or stimulan* or Narcotic*) near/3 (abus* or us* or misus* or addict* or depend*):ti,ab,kw	61546
s47	MeSH descriptor Substance Abuse, Intravenous explode all trees	310
s48	(#23 OR #47)	3660
s49	(#7 OR #16)	1869
s50	(#48 AND #49)	67
s51	(#45 AND #49)	393
s52	(#46 AND #49)	181
S53	(#50 OR #51 OR #52)	399
S54	(#50 OR #51 OR #52) - Limited to 2015-current	4

Embase

Searched January 2011.

Update September 2011, July 2013, August 2015 and final update September 2016.

Search number	Term	totals
s1	(narcotics anonymous or Cocaine Anonymous or Crystal Meth Anonymous or Pills Anonymous or Marijuana Anonymous or Heroin Anonymous).ab,kw,ti.	117
s2	((Self-Help adj1 group*) or (Support* adj1 group*)).ab,ti.	9266
s3	(twelve-step* or 12-step*).ab,ti.	1334
s4	Self-Help Groups/	10817
s5	(Recover* adj1 group*).ab,ti.	1821
s6	(Minnesota adj3 (model* or program* or treatment*or Rehab* or cure* or therap* or detox* or recover* or intervent* or method*)).ab,ti.	576
s7	TSF.ab,ti.	1044

s8	1 or 2 or 3 or 4 or 5 or 6 or 7	22601
s9	Amphetamine-Related Disorders/ or Cocaine-Related Disorders/ or Marijuana Abuse/ or opioid-Related Disorders/ or Substance Abuse, Intravenous/ or Substance Withdrawal Syndrome/ or heroin Dependence/	116800
s10	Opium/ or Heroin/ or exp Cannabinoids/ or Cannabis/ or Marijuana Smoking/ or exp cocaine/ or methamphetamine/ or Amphetamine/ or Designer Drugs/	158507
s11	(Drug* or Substance* or Stimulan* or Narcotic* or Opium or Heroin or Crack or Cocaine* or (Methamphetamine* or Amphetamin*) or Ecstasy or Fantasy or (Cannabis or Marijuana or Hashish) or (Designer adj1 drug*)).ab,ti.	1691746
s12	((substance* or drug* or stimulan* or Narcotic*) adj3 (abus* or us* or misus* or addict* or depend*)).ab,ti.	268568
s13	9 or 10 or 11	1781908
s14	8 and 13	2293
s15	limit 14 to humans	1978
s16	(narcotics anonymous or Cocaine Anonymous or Crystal Meth Anonymous or Pills Anonymous or Marijuana Anonymous or Heroin Anonymous).ab,kw,ti.	117
s17	((Self-Help adj1 group*) or (Support* adj1 group*)).ab,ti.	9266
s18	(twelve-step* or 12-step*).ab,ti.	1334
s19	Self-Help Groups/	10817
s20	(Recover* adj1 group*).ab,ti.	1821
s21	(Minnesota adj3 (model* or program* or treatment*or Rehab* or cure* or therap* or detox* or recover* or intervent* or method*)).ab,ti.	576
s22	TSF.ab,ti.	1044
s23	16 or 17 or 18 or 19 or 20 or 21 or 22	22601
s24	Amphetamine-Related Disorders/ or Cocaine-Related Disorders/ or Marijuana Abuse/ or opioid-Related Disorders/ or Substance Abuse,	116800

	Intravenous/ or Substance Withdrawal Syndrome/ or heroin Dependence/	
s25	Opium/ or Heroin/ or exp Cannabinoids/ or Cannabis/ or Marijuana Smoking/ or exp cocaine/ or methamphetamine/ or Amphetamine/ or Designer Drugs/	158507
s26	(Drug* or Substance* or Stimulan* or Narcotic* or Opium or Heroin or Crack or Cocaine* or (Methamphetamine* or Amphetamin*) or Ecstasy or Fantasy or (Cannabis or Marijuana or Hashish) or (Designer adj1 drug*)).ab,ti.	1691746
s27	((substance* or drug* or stimulan* or Narcotic*) adj3 (abus* or us* or misus* or addict* or depend*)).ab,ti.	268568
s28	24 or 25 or 26	1781908
s29	23 and 28	2293
s30	limit 29 to humans	1978
s31	limit 30 to yr="2015 -Current"	214
s32	opiate*.ti,ab.	29806
s33	8 and 32	68
s34	limit 33 to human	66
s35	31 or 34	277

LIBRIS

Searched July 2013, August 2015 and final update September 2016.

Search	Term	Totals
s1	(narcotics anonymous) or (cocaine anonymous*) or (crystal meth anonymous*) or (pills anonymous*) or (marijuana anonymous*) eller (heroin anonymous*) or (12 step) or (hjälpgrupp*) eller (selfhelp group?) or (minnesota model*) or (minnesota program*) or (minnesota behandling?) or (tolvstegsbehandling*)	11

s2	(narcotics anonymous) or (cocaine anonymous*) or (crystal meth anonymous*) or (pills anonymous*) or (marijuana anonymous*) eller (heroin anonymous*) or (12 step) or (hjälpgrupp*) eller (selfhelp group?) or (minnesota model*) or (minnesota program*) or (minnesota behandling?) or (tolvstegsbehandling*) - Limited to 2015-current	0

Medline

Searched January 2010.

Update September 2011, July 2013, August 2015 and final update September 2016.

Search Number	Term	Totals
1	narcotics anonymous.ab,kw,ti.	63
2	Cocaine Anonymous.ab,kw,ti.	3
3	Crystal Meth Anonymous.ab,kw,ti.	2
4	Pills Anonymous.ab,kw,ti.	1
5	Marijuana Anonymous.ab,kw,ti.	0
6	Heroin Anonymous.ab,kw,ti.	0
7	1 or 2 or 3 or 4 or 5 or 6	66
8	(Self-Help adj1 group*).ab,ti.	1175
9	(Support* adj1 group*).ab,ti.	4699
10	twelve-step*.ab,ti.	159
11	12-step*.ab,ti.	613
12	Self-Help Groups/	6943
13	(Recover* adj1 group*).ab,ti.	1151
14	(Minnesota and (model* or program* or treatment*or Rehab* or cure* or therap* or detox* or recover*)).ab,ti.	3024
15	TSF.ab,ti.	715
16	8 or 9 or 10 or 11 or 12 or 13 or 14 or 15	16194
17	Amphetamine-Related Disorders/	1547
18	Cocaine-Related Disorders/	5105
19	Marijuana Abuse/	3417
20	Opioid-Related Disorders/	6644
21	Substance Abuse, Intravenous/	10829
22	Substance Withdrawal Syndrome/	17575

23	Heroin Dependence/	7294
24	17 or 18 or 19 or 20 or 21 or 22 or 23	47243
25	Opium/	1678
26	Heroin/	4497
27	exp Cannabinoids/ or Cannabis/	12874
28	Marijuana Smoking/	1848
29	exp cocaine/	20516
30	Methamphetamine/	5733
31	Amphetamine/	10679
32	Designer Drugs/	539
33	Drug*.ab,ti.	863790
34	Substance*.ab,ti.	180252
35	Stimulan*.ab,ti.	16686
36	Narcotic*.ab,ti.	10913
37	Opium.ab,ti.	1285
38	Heroin.ab,ti.	8879
39	Crack.ab,ti.	3436
40	Cocaine*.ab,ti.	25231
41	(Methamphetamine* or Amphetamin*).ab,ti.	22897
42	Ecstasy.ab,ti.	2317
43	Fantasy.ab,ti.	1895
44	(Cannabis or Marijuana or Hashish).ab,ti.	11762
45	(Designer adj1 drug*).ab,ti.	425
46	25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45	1081525
47	(substance adj1 abus*).ab,ti.	15078
48	(substance adj1 us*).ab,ti.	12077
49	(substance adj1 addict*).ab,ti.	184
50	(drug adj1 us*).ab,ti.	37875
51	(drug adj1 usage).ab,ti.	870
52	(drug adj1 abus*).ab,ti.	14242

53	(drug adj1 addict*).ab,ti.	7340
54	(drug adj1 depend*).ab,ti.	5031
55	47 or 48 or 49 or 50 or 51 or 52 or 53 or 54	79271
56	7 or 16	16211
57	24 and 56	224
58	46 and 56	1532
59	55 and 56	811
60	57 or 58 or 59	1579
61	limit 60 to humans	1474
62	limit 61 to yr="2015 -Current"	78
63	opiate*.ti,ab	21965
64	56 and 63	48
65	limit 64 to humans	46
66	62 or 65	124

PsycINFO

Searched December 2010.

Update September 2011, July 2013, August 2015 and final update September 2016.

Search	Terms	Limiters	Total
S57	S54 OR S55	**Limiters** - Date of Publication: 20150901-Current **Search modes** - Boolean/Phrase	(312)
S56	S54 OR S55	**Search modes** - Boolean/Phrase	(278,755)
S55	S51 AND S53	**Search modes** - Boolean/Phrase	(1,892)
S54	S26 OR S53	**Search modes** - Boolean/Phrase	(278,755)
S53	S19 OR S52	**Search modes** - Boolean/Phrase	(16,660)
S52	TI "TSF" OR AB TSF	**Search modes** - Boolean/Phrase	(130)
S51	S32 OR S38 OR S44 OR S50	**Search modes** - Boolean/Phrase	(119,611)
S50	S45 OR S46 OR S47 OR S48 OR S49	**Search modes** - Boolean/Phrase	(1,290)
S49	TI Narcotic* n3 depend* OR AB Narcotic* n3 depend*	**Search modes** - Boolean/Phrase	(117)
S48	TI Narcotic* n3 addict* OR AB Narcotic* n3 addict*	**Search modes** - Boolean/Phrase	(701)

S47	TI Narcotic* n3 misus* OR AB Narcotic* n3 misus*	**Search modes** - Boolean/Phrase	(14)
S46	TI Narcotic* n3 us* OR AB Narcotic* n3 us*	**Search modes** - Boolean/Phrase	(495)
S45	TI Narcotic* n3 abus* OR AB Narcotic* n3 abus*	**Search modes** - Boolean/Phrase	(125)
S44	S39 OR S40 OR S41 OR S42 OR S43	**Search modes** - Boolean/Phrase	(2,612)
S43	TI stimulan* n3 depend* OR AB stimulan* n3 depend*	**Search modes** - Boolean/Phrase	(322)
S42	TI stimulan* n3 addict* OR AB stimulan* n3 addict*	**Search modes** - Boolean/Phrase	(163)
S41	TI stimulan* n3 misus* OR AB stimulan* n3 misus*	**Search modes** - Boolean/Phrase	(96)
S40	TI stimulan* n3 us* OR AB stimulan* n3 us*	**Search modes** - Boolean/Phrase	(1,895)
S39	TI stimulan* n3 abus* OR AB stimulan* n3 abus*	**Search modes** - Boolean/Phrase	(537)
S38	S33 OR S34 OR S35 OR S36 OR S37	**Search modes** - Boolean/Phrase	(73,777)
S37	TI drug* n3 depend* OR AB drug* n3 depend*	**Search modes** - Boolean/Phrase	(7,183)
S36	TI drug* n3 addict* OR AB drug* n3 addict*	**Search modes** - Boolean/Phrase	(9,069)
S35	TI drug* n3 misus* OR AB drug* n3 misus*	**Search modes** - Boolean/Phrase	(1,703)
S34	TI drug* n3 us* OR AB drug* n3 us*	**Search modes** - Boolean/Phrase	(50,706)
S33	TI drug* n3 abus* OR AB drug* n3 abus*	**Search modes** - Boolean/Phrase	(19,282)
S32	S27 OR S28 OR S29 OR S30 OR S31	**Search modes** - Boolean/Phrase	(59,214)
S31	TI substance* n3 misus* OR AB substance* n3 misus*	**Search modes** - Boolean/Phrase	(2,790)
S30	TI substance* n3 depend* OR AB substance* n3 depend*	**Search modes** - Boolean/Phrase	(4,561)
S29	TI substance* n3 addict* OR AB substance* n3 addict*	**Search modes** - Boolean/Phrase	(1,981)
S28	TI substance* n3 us* OR AB substance* n3 us*	**Search modes** - Boolean/Phrase	(32,012)
S27	TI substance* n3 abus* OR AB substance* n3 abus*	**Search modes** - Boolean/Phrase	(31,439)
S26	S20 OR S21 OR S22 OR S23 OR S24 OR S25	**Search modes** - Boolean/Phrase	(264,505)

S25	TI Designer n1 drug* OR AB Designer n1 drug*	**Search modes** - Boolean/Phrase	(213)
S24	TI ((Opium or Heroin or Crack or Cocaine* or Methamphetamine* or Amphetamin*or Ecstasy or Fantasy or Cannabis or Marijuana or Hashish)) OR AB ((Opium or Heroin or Crack or Cocaine* or Methamphetamine* or Amphetamin*or Ecstasy or Fantasy or Cannabis or Marijuana or Hashish))	**Search modes** - Boolean/Phrase	(54,540)
S23	TI Narcotic* OR AB Narcotic*	**Search modes** - Boolean/Phrase	(3,060)
S22	TI Stimulan* OR AB Stimulan*	**Search modes** - Boolean/Phrase	(9,366)
S21	TI Substance* OR AB Substance*	**Search modes** - Boolean/Phrase	(75,860)
S20	TI Drug* OR AB Drug*	**Search modes** - Boolean/Phrase	(179,397)
S19	S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S18	**Search modes** - Boolean/Phrase	(16,614)
S18	S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17	**Search modes** - Boolean/Phrase	(459)
S17	TI Minnesota n3 method* OR AB Minnesota n3 method*	**Search modes** - Boolean/Phrase	(83)
S16	TI Minnesota n3 intervent* OR AB Minnesota n3 intervent*	**Search modes** - Boolean/Phrase	(22)
S15	TI Minnesota n3 recover* OR AB Minnesota n3 recover*	**Search modes** - Boolean/Phrase	(2)
S14	TI Minnesota n3 detox* OR AB Minnesota n3 detox*	**Search modes** - Boolean/Phrase	(0)
S13	TI Minnesota n3 therap* OR AB Minnesota n3 therap*	**Search modes** - Boolean/Phrase	(22)
S12	TI Minnesota n3 cure* OR AB Minnesota n3 cure*	**Search modes** - Boolean/Phrase	(0)
S11	TI Minnesota n3 rehab* OR AB Minnesota n3 rehab*	**Search modes** - Boolean/Phrase	(19)
S10	TI Minnesota n3 treatment* OR AB Minnesota n3 treatment*	**Search modes** - Boolean/Phrase	(79)
S9	TI Minnesota n3 program* OR AB Minnesota n3 program*	**Search modes** - Boolean/Phrase	(216)
S8	TI Minnesota n3 model* OR AB Minnesota n3 model*	**Search modes** - Boolean/Phrase	(95)
S7	TI (Recover* n1 group*) OR AB (Recover* n1 group*)	**Search modes** - Boolean/Phrase	(698)

S6	SU "SELF-help groups" OR SU "SELF-help groups for substance abusers"	**Search modes** - Boolean/Phrase	(3,025)
S5	TI (twelve-step* or twelve n1 step*) OR AB (twelve-step* or twelve n1 step*)	**Search modes** - Boolean/Phrase	(1,952)
S4	TI (12-step* OR 12 n1 step*) OR AB (12-step* OR 12 n1 step*)	**Search modes** - Boolean/Phrase	(1,952)
S3	TI Support* n1 group* OR AB Support* n1 group*	**Search modes** - Boolean/Phrase	(10,561)
S2	TI Self-Help N1 group* OR AB Self-Help N1 group*	**Search modes** - Boolean/Phrase	(1,905)
S1	TI ((narcotics anonymous or Cocaine Anonymous or Crystal Meth Anonymous or Pills Anonymous or Marijuana Anonymous or Heroin Anonymous)) OR AB ((narcotics anonymous or Cocaine Anonymous or Crystal Meth Anonymous or Pills Anonymous or Marijuana Anonymous or Heroin Anonymous))	**Search modes** - Boolean/Phrase	(237)

Science Citation Index

Searched December 2010.

Update September 2011, July 2013, and final update September 2016.

In the 2016 update, Science Citation Index and Social Science Citation Index were searched together. Search #47 is the total result from both Science Citation Index and Social Science Citation Index in the 2016 update.

Search number	Term	Totals
s1	TS=("narcotics anonymous") OR TI=("narcotics anonymous") Databases=SCI-EXPANDED Timespan=All Years	32
s2	TS=("cocaine anonymous") OR TI=("cocaine anonymous") Databases=SCI-EXPANDED Timespan=All Years	3
s3	TS=("Crystal Meth Anonymous") OR TI=("Crystal Meth Anonymous") Databases=SCI-EXPANDED Timespan=All Years	2
s4	TS=("Pills Anonymous") OR TI=("Pills Anonymous") Databases=SCI-EXPANDED Timespan=All Years	1

s5	TS=("marijuana anonymous") OR TI=("marijuana anonymous") Databases=SCI-EXPANDED Timespan=All Years	0
s6	TS="heroin anonymous" OR TI="heroin anonymous" Databases=SCI-EXPANDED Timespan=All Years	0
s7	#6 OR #5 OR #4 OR #3 OR #2 OR #1 Databases=SCI-EXPANDED Timespan=All Years	34
s8	TS=(self-help SAME group*) OR TI=(self-help SAME group*) Databases=SCI-EXPANDED Timespan=All Years	777
s9	TS=(support* SAME group*) OR TI=(support* SAME group*) Databases=SCI-EXPANDED Timespan=All Years	26,736
s10	TS=twelve-step* OR TI=twelve-step* Databases=SCI-EXPANDED Timespan=All Years	131
s11	TS=12-step* OR TI=12-step* Databases=SCI-EXPANDED Timespan=All Years	547
s12	TS=("self-help group*") OR TI=("self-help group*") Databases=SCI-EXPANDED Timespan=All Years	598
s13	TS=(recover* SAME group*) OR TI=(recover* SAME group*) Databases=SCI-EXPANDED Timespan=All Years	15,293
s14	TS=(Minnesota SAME (model* or program* or treatment*or Rehab* or cure* or therap* or detox* or recover* or intervent* or method*)) OR TI=(Minnesota SAME (model* or program* or treatment*or Rehab* or cure* or therap* or detox* or recover* or intervent* or method*)) Databases=SCI-EXPANDED Timespan=All Years	1,768
s15	TS=TSF OR TI=TSF Databases=SCI-EXPANDED Timespan=All Years	657
s16	#15 OR #14 OR #13 OR #12 OR #11 OR #10 OR #9 OR #8 Databases=SCI-EXPANDED Timespan=All Years	45,216
s17	TS=(amphetamine-related SAME disorder*) OR TI=(amphetamine-related SAME disorder*)Databases=SCI-EXPANDED Timespan=All Years	7
s18	TS=(Cocaine-Related SAME Disorder*) OR TI=(Cocaine-Related SAME Disorder*) Databases=SCI-EXPANDED Timespan=All Years	30
s19	TS=(Marijuana SAME Abuse) OR TI=(Marijuana SAME Abuse) Databases=SCI-EXPANDED Timespan=All Years	258

s20	TS=(opioid-related SAME disorder*) OR TI=(opioid-related SAME disorders) Databases=SCI-EXPANDED Timespan=All Years	52
s21	TS=((substance SAME abuse) AND Intravenous) OR TI=((substance SAME abuse) AND Intravenous) Databases=SCI-EXPANDED Timespan=All Years	342
s22	TS=(substance SAME withdrawal SAME syndrome) OR TI=(substance SAME withdrawal SAME syndrome) Databases=SCI-EXPANDED Timespan=All Years	74
s23	TS=(heroin SAME dependence) OR TI=(heroin SAME dependence) Databases=SCI-EXPANDED Timespan=All Years	541
s24	#23 OR #22 OR #21 OR #20 OR #19 OR #18 OR #17 Databases=SCI-EXPANDED Timespan=All Years	1,273
s25	TS=opium OR TI=opium Databases=SCI-EXPANDED Timespan=All Years	1,652
s26	TS=heroin OR TI=heroin Databases=SCI-EXPANDED Timespan=All Years	8,746
s27	TS=cannabi* OR TI=cannabi* Databases=SCI-EXPANDED Timespan=All Years	18,094
s28	TS=(marijuana SAME smoking) OR TI=(marijuana SAME smoking) Databases=SCI-EXPANDED Timespan=All Years	524
s29	TS=cocaine* OR TI=cocaine* Databases=SCI-EXPANDED Timespan=All Years	30,927
s30	TS=methamphetamine* OR TI=methamphetamine* Databases=SCI-EXPANDED Timespan=All Years	7,149
s31	TS= amphetamine* OR TI= amphetamine* Databases=SCI-EXPANDED Timespan=All Years	22,692
s32	TS=(designer SAME drug*) OR TI=(designer SAME drug*) Databases=SCI-EXPANDED Timespan=All Years	656
s33	TS=drug* OR TI=drug* Databases=SCI-EXPANDED Timespan=All Years	>100,000
s34	TS=substance* OR TI=substance* Databases=SCI-EXPANDED Timespan=All Years	>100,000
s35	TS=stimulan* OR TI=stimulan* Databases=SCI-EXPANDED Timespan=All Years	14,591
s36	TS=narcotic* OR TI=narcotic*Databases=SCI-EXPANDED Timespan=All Years	8,407

s37	TS=crack OR TI=crack Databases=SCI-EXPANDED Timespan=All Years	72,462
s38	TS=ectasy OR TI=ectasy Databases=SCI-EXPANDED Timespan=All Years	24
s39	TS=fantasy OR TI=fantasy Databases=SCI-EXPANDED Timespan=All Years	1,159
s40	TS=(marijuana OR hashish) OR TI=(marijuana OR hashish) Databases=SCI-EXPANDED Timespan=All Years	5,524
s41	#40 OR #39 OR #38 OR #37 OR #36 OR #35 OR #34 OR #33 OR #32 OR #31 OR #30 OR #29 OR #28 OR #27 OR #26 OR #25 Databases=SCI-EXPANDED Timespan=All Years	>100,000
s42	TS=((substance* or drug* or stimulan* or Narcotic*) SAME (abus* or use* or misus* or addict* or depend*)) OR TI=((substance* or drug* or stimulan* or Narcotic*) SAME (abus* or use* or misus* or addict* or depend*)) Databases=SCI-EXPANDED Timespan=All Years	>100,000
s43	#16 OR #7 Databases=SCI-EXPANDED Timespan=All Years	45,23
s44	#43 AND #24 Databases=SCI-EXPANDED Timespan=All Years	12
s45	#43 AND #41 Databases=SCI-EXPANDED Timespan=All Years	3,192
s46	#43 AND #42 Databases=SCI-EXPANDED Timespan=All Years	1,049
s47	#46 OR #45 OR #44 Databases=SCI-EXPANDED, SSCI Timespan=2015-2016	1.176

Social Science Citation Index

Searched January 2011

Update September 2011, July 2013, September 2015 and final update September 2016.

In the 2016 update, Science Citation Index and Social Science Citation Index were searched together. See the searches for Science Citation Index for 2016 results.

Search number	Term	Totals
1	Topic=(("narcotics anonymous")) OR Title=(("narcotics anonymous")) Databases=SSCI Timespan=All years	16
2	Topic=(("narcotics anonymous")) OR Title=(("narcotics anonymous")) Databases=SSCI Timespan=All years	16

3	Topic=(("cocaine anonymous")) OR Title=(("cocaine anonymous")) Databases=SSCI Timespan=All years	0
4	Topic=(("Crystal Meth Anonymous")) OR Title=(("Crystal Meth Anonymous")) Databases=SSCI Timespan=All years	0
5	Topic=("marijuana anonymous") OR Title=("marijuana anonymous") Databases=SSCI Timespan=All years	0
6	Topic=("heroin anonymous") OR Title=("heroin anonymous") Databases=SSCI Timespan=All years	0
7	#6 OR #5 OR #4 OR #3 OR #2 OR #1 Databases=SSCI Timespan=All years Databases=SSCI Timespan=All years	16
8	Topic=((self-help SAME group*)) OR Title=((self-help SAME group*)) Databases=SSCI Timespan=All years	341
9	Topic=((support* SAME group*)) OR Title=((support* SAME group*)) Databases=SSCI Timespan=All years	11,754
10	Topic=(twelve-step*) OR Title=(twelve-step*) Databases=SSCI Timespan=All years	24
11	Topic=(12-step*) OR Title=(12-step*) Databases=SSCI Timespan=All years	89
12	Topic=(("self-help group*")) OR Title=(("self-help group*")) Databases=SSCI Timespan=All years	115
13	Topic=((recover* SAME group*)) OR Title=((recover* SAME group*)) Databases=SSCI Timespan=All years	1,284
14	Topic=((Minnesota SAME (model* or program* or treatment*or Rehab* or cure* or therap* or detox* or recover* or intervent* or method*))) OR Title=((Minnesota SAME (model* or program* or treatment*or Rehab* or cure* or therap* or detox* or recover* or intervent* or method*))) Databases=SSCI Timespan=All years	311
15	Topic=(TSF) OR Title=(TSF) Databases=SSCI Timespan=All years	19
16	#15 OR #14 OR #13 OR #12 OR #11 OR #10 OR #9 OR #8 Databases=SSCI Timespan=All years	13,251
17	Topic=((amphetamine-related SAME disorder*)) OR Title=((amphetamine-related SAME disorder*)) Databases=SSCI Timespan=All years	1
18	Topic=((Cocaine-Related SAME Disorder*)) OR Title=((Cocaine-Related SAME Disorder*)) Databases=SSCI Timespan=All years	5

19	Topic=(Marijuana SAME Abuse) OR Title=(Marijuana SAME Abuse) Databases=SSCI Timespan=All years	284
20	Topic=(opioid-related SAME disorder*) OR Title=(opioid-related SAME disorder*) Databases=SSCI Timespan=All years	25
21	Topic=((substance SAME abuse) AND Intravenous) OR Title=((substance SAME abuse) AND Intravenous) Databases=SSCI Timespan=All years	28
22	Topic=((substance SAME withdrawal SAME syndrome)) OR Title=((substance SAME withdrawal SAME syndrome)) Databases=SSCI Timespan=All years	30
23	Topic=(heroin SAME dependen*) OR Title=(heroin SAME dependen*) Databases=SSCI Timespan=All years	334
24	#23 OR #22 OR #21 OR #20 OR #19 OR #18 OR #17 Databases=SSCI Timespan=All years	672
25	Topic=(opium) OR Title=(opium) Databases=SSCI Timespan=All years	49
26	Topic=(heroin) OR Title=(heroin) Databases=SSCI Timespan=All years	686
27	Topic=(cannabi*) OR Title=(cannabi*) Databases=SSCI Timespan=All years	1,15
28	Topic=(marijuana SAME smoking) OR Title=(marijuana SAME smoking) Databases=SSCI Timespan=All years	224
29	Topic=(cocain*) OR Title=(cocain*) Databases=SSCI Timespan=All years	1,095
30	Topic=(methamphetamine*) OR Title=(methamphetamine*) Databases=SSCI Timespan=All years	387
31	Topic=(amphetamine*) OR Title=(amphetamine*) Databases=SSCI Timespan=All years	369
32	Topic=(designer SAME drug*) OR Title=(designer SAME drug*) Databases=SSCI Timespan=All years	31
33	Topic=(drug*) OR Title=(drug*) Databases=SSCI Timespan=All years	13,241
34	Topic=(substance*) OR Title=(substance*) Databases=SSCI Timespan=All years	7,827
35	Topic=(stimulan*) OR Title=(stimulan*) Databases=SSCI Timespan=All years	524
36	Topic=(narcotic*) OR Title=(narcotic*) Databases=SSCI Timespan=All years	132

37	Topic=(crack) OR Title=(crack) Databases=SSCI Timespan=All years	307
38	Topic=(ecstasy) OR Title=(ecstasy) Databases=SSCI Timespan=All years	147
39	Topic=(fantasy) OR Title=(fantasy) Databases=SSCI Timespan=All years	449
40	Topic=(marijuana OR hashish) OR Title=(marijuana OR hashish) Databases=SSCI Timespan=All years	877
41	#40 OR #39 OR #38 OR #37 OR #36 OR #35 OR #34 OR #33 OR #32 OR #31 OR #30 OR #29 OR #28 OR #27 OR #26 OR #25 Databases=SSCI Timespan=All years	19,44
42	Topic=((substance* or drug* or stimulan* or Narcotic*) SAME (abus* or use* or misus* or addict* or depend*)) OR Title=((substance* or drug* or stimulan* or Narcotic*) SAME (abus* or use* or misus* or addict* or depend*)) Databases=SSCI Timespan=All years	14,016
43	#16 OR #7 Databases=SSCI Timespan=All years	13,256
44	#43 AND #42 Databases=SSCI Timespan=All years	782
45	#43 AND #41 Databases=SSCI Timespan=All years	930
46	#43 AND #24 Databases=SSCI Timespan=All years	27
47	#46 OR #45 OR #44 Databases=SSCI Timespan=All years	930
48	Topic=(opiate*) AND Title=(opiate*) Databases=SSCI Timespan=All years	1,758
49	#48 AND #43 Databases=SSCI Timespan=All years	5
50	#49 OR #47 Databases=SSCI Timespan=All years	930

SocIndex

Searched December 2010.

Update September 2011, July 2013, August 2015 and final update September 2016.

Search number	Term	Totals
s1	TI (narcotics anonymous or Cocaine Anonymous or Crystal Meth Anonymous or Pills Anonymous or Marijuana Anonymous or Heroin Anonymous) or AB (narcotics anonymous or Cocaine Anonymous or Crystal Meth	122

	Anonymous or Pills Anonymous or Marijuana Anonymous or Heroin Anonymous)	
s2	TI Self-Help N1 group* or AB Self-Help N1 group*	824
s3	TI Support* n1 group* or AB Support* n1 group*	3194
s4	TI (12-step* or 12 n1 step*) or AB (12-step* or 12 n1 step*)	646
s5	TI (twelve-step* or twelve n1 step*) or AB (twelve-step* or twelve n1 step*)	646
s6	DE "SELF-help groups" OR DE "SELF-help groups for substance abusers"	1154
s7	TI (Recover* n1 group*) and AB (Recover* n1 group*)	1
s8	TI Minnesota n3 model* or AB Minnesota n3 model*	66
s9	TI Minnesota n3 program* or AB Minnesota n3 program*	232
s10	TI Minnesota n3 treatment* or AB Minnesota n3 treatment*	48
s11	TI Minnesota n3 rehab* or AB Minnesota n3 rehab*	11
s12	TI Minnesota n3 cure* or AB Minnesota n3 cure*	0
s13	TI Minnesota n3 therap* or AB Minnesota n3 therap*	14
s14	TI Minnesota n3 detox* or AB Minnesota n3 detox*	0
s15	TI Minnesota n3 recover* or AB Minnesota n3 recover*	2
s16	TI Minnesota n3 intervent* or AB Minnesota n3 intervent*	7
s17	TI Minnesota n3 method* or AB Minnesota n3 method*	30
s18	S8 or S9 or S10 or S11 or S12 or S13 or S14 or S15 or S16 or S17	366
s19	S1 or S2 or S3 or S4 or S5 or S6 or S7 or S18	5354
s20	TI Drug* or AB Drug*	67211
s21	TI Substance* or AB Substance*	23945
s22	TI Stimulan* or AB Stimulan*	1091
s23	TI Narcotic* or AB Narcotic*	4104
s24	TI (Opium or Heroin or Crack or Cocaine* or Methamphetamine* or Amphetamin*or Ecstasy or Fantasy or Cannabis or Marijuana or Hashish) or AB (Opium or Heroin or Crack or Cocaine* or Methamphetamine* or Amphetamin*or Ecstasy or Fantasy or Cannabis or Marijuana or Hashish)	18378
s25	TI Designer n1 drug* or AB Designer n1 drug*	88
s26	S20 or S21 or S22 or S23 or S24 or S25	92592
s27	TI substance* n3 abus* or AB substance* n3 abus*	13074

s28	TI substance* n3 us* or AB substance* n3 us*	7400
s29	TI substance* n3 addict* or AB substance* n3 addict*	447
s30	TI substance* n3 depend* or AB substance* n3 depend*	932
s31	TI substance* n3 misus* or AB substance* n3 misus*	711
s32	S27 or S28 or S29 or S30 or S31	19148
s33	TI drug* n3 abus* or AB drug* n3 abus*	14035
s34	TI drug* n3 us* or AB drug* n3 us*	21927
s35	TI drug* n3 misus* or AB drug* n3 misus*	939
s36	TI drug* n3 addict* or AB drug* n3 addict*	3856
s37	TI drug* n3 depend* or AB drug* n3 depend*	2412
s38	S33 or S34 or S35 or S36 or S37	35750
s39	TI stimulan* n3 abus* or AB stimulan* n3 abus*	96
s40	TI stimulan* n3 us* or AB stimulan* n3 us*	258
s41	TI stimulan* n3 misus* or AB stimulan* n3 misus*	8
s42	TI stimulan* n3 addict* or AB stimulan* n3 addict*	25
s43	TI stimulan* n3 depend* or AB stimulan* n3 depend*	35
s44	S39 or S40 or S41 or S42 or S43	365
s45	TI Narcotic* n3 abus* or AB Narcotic* n3 abus*	296
s46	TI Narcotic* n3 us* or AB Narcotic* n3 us*	441
s47	TI Narcotic* n3 misus* or AB Narcotic* n3 misus*	9
s48	TI Narcotic* n3 addict* or AB Narcotic* n3 addict*	584
s49	TI Narcotic* n3 depend* or AB Narcotic* n3 depend*	37
s50	S45 or S46 or S47 or S48 or S49	1232
s51	S32 or S38 or S44 or S50	50035
s52	TI TSF or AB tsf	27
s53	(S19 or S52)	5356
s54	S26 and S53	933
s55	S51 and S53	710
s56	(S54 or S55)	933
S57	(S54 or S55) - Date of Publication:20150901-20160901	64

Dissertation Abstracts

Searched December 2010.

Update September 2011, July 2013, August 2015 and final update September 2016.

Search	Terms	Total
S4	(((narcotics anonymous) OR (Cocaine Anonymous) OR (Crystal Meth Anonymous) OR (Crystal Meth Anonymous) OR (Pills Anonymous) OR (Marijuana Anonymous) OR (Heroin Anonymous)) OR (Self-Help NEAR/1 group) OR (Support NEAR/1 group) OR (twelve-step*) OR (12-step*) OR (Recover NEAR/1 group) OR TSF OR (Minnesota AND (model* OR program* OR treatment*or Rehab* OR cure* OR therap* OR detox* OR recover* OR intervent* OR method*))) AND ((su(Heroin) OR su(Marijuana) OR su(cannabis) OR su(cocaine) OR su(Methamphetamine) OR su(Amphetamine) OR drug* OR Substance* OR Stimulan* OR Narcotic* OR Opium OR Heroin OR Crack OR Cocaine* OR Ecstasy OR Fantasy OR (designer NEAR/1 drug*)) OR ((substance* OR drug* OR stimulan* OR Narcotic*) AND (substance* OR drug* OR stimulan* OR Narcotic*) AND (abus* OR us* OR misus* OR addict* OR depend*)) - Limited to 2015-current	49
S3	(su(Heroin) OR su(Marijuana) OR su(cannabis) OR su(cocaine) OR su(Methamphetamine) OR su(Amphetamine) OR drug* OR Substance* OR Stimulan* OR Narcotic* OR Opium OR Heroin OR Crack OR Cocaine* OR Ecstasy OR Fantasy OR (designer NEAR/1 drug*)) OR ((substance* OR drug* OR stimulan* OR Narcotic*) AND (abus* OR us* OR misus* OR addict* OR depend*))	10.499
S2	((narcotics anonymous) OR (Cocaine Anonymous) OR (Crystal Meth Anonymous) OR (Crystal Meth Anonymous) OR (Pills Anonymous) OR (Marijuana Anonymous) OR (Heroin Anonymous)) OR (Self-Help NEAR/1 group) OR (Support NEAR/1 group) OR (twelve-step*) OR (12-step*) OR (Recover NEAR/1 group) OR TSF OR (Minnesota AND (model* OR program* OR treatment*or Rehab* OR cure* OR therap* OR detox* OR recover* OR intervent* OR method*))	145
S1	((substance* OR drug* OR stimulan* OR Narcotic*) AND (abus* OR us* OR misus* OR addict* OR depend*))Limits applied	2033

Sociological Abstracts

Searched December 2010.

Update September 2011, July 2013, August 2015 and final update September 2016.

Search	Terms	Total
S4	(("narcotics anonymous" OR "Cocaine Anonymous" OR "Crystal Meth Anonymous" OR "Pills Anonymous" OR "Marijuana Anonymous" OR "Heroin Anonymous") OR (Self-Help NEAR/1 group) OR (Support NEAR/1 group) OR ((12-step*) OR (twelve-step*)) OR (Recover NEAR/1 group) OR "TSF" OR ((Minnesota) AND (model* OR program* OR treatment*or Rehab* OR cure* OR therap* OR detox* OR recover* OR intervent* OR method*))) AND ((su(heroin) OR su((marijuana OR cannabis)) OR su((cocaine OR Methamphetamine)) OR su(Amphetamine)) OR ((drug*) OR ((substance*) OR (stimulan*)) OR ((narcotic*) OR (opium)) OR ((heroin*) OR (crack*)) OR ((cocaine*) OR (ecstasy*)) OR ((fantas*) OR (designer NEAR/1 drug*))) OR ((substance* OR drug* OR stimulan* OR Narcotic*) AND (abus* OR us* OR misus* OR addict* OR depend*))) Limited to 2015 - current.	7

S3	(su(heroin) OR su((marijuana OR cannabis)) OR su((cocaine OR Methamphetamine)) OR su(Amphetamine)) OR ((drug*) OR ((substance*) OR (stimulan*)) OR ((narcotic*) OR (opium)) OR ((heroin*) OR (crack*)) OR ((cocaine*) OR (ecstasy*)) OR ((fantas*) OR (designer NEAR/1 drug*))) OR ((substance* OR drug* OR stimulan* OR Narcotic*) AND (abus* OR us* OR misus* OR addict* OR depend*))	577
S2	("narcotics anonymous" OR "Cocaine Anonymous" OR "Crystal Meth Anonymous" OR "Pills Anonymous" OR "Marijuana Anonymous" OR "Heroin Anonymous") OR (Self-Help NEAR/1 group) OR (Support NEAR/1 group) OR ((12-step*) OR (twelve-step*)) OR (Recover NEAR/1 group) OR "TSF" OR ((Minnesota) AND (model* OR program* OR treatment*or Rehab* OR cure* OR therap* OR detox* OR recover* OR intervent* OR method*))	52
S1	((substance* OR drug* OR stimulan* OR Narcotic*) AND (abus* OR us* OR misus* OR addict* OR depend*))	294

SveMed+

Searched December 2010.

Update September 2011 and July 2013 and final update September 2016.

Search	Terms	Total
1	narcotics anonymous	60
2	Cocaine Anonymous	1
3	Crystal Meth Anonymous	0
4	Pills Anonymous	4
5	Marijuana Anonymous	1
6	Heroin Anonymous	3
7	#1 OR #2 OR #3 OR #4 OR #5 OR #6	68
8	Self-Help group*	179
9	Support* group*	253
10	twelve-step*	0
11	12-step*	0
12	12-step	16
13	Self-Help Groups	187
14	recovery group	27
15	minnesota model	4
16	minnesota program	0
17	minnesota treatment	11
18	minnesota rehab	0
19	minnesota cure	0
20	minnesota therapy	9
21	minnesota detox	0
22	minnesota recover	0
23	#8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22	407
24	Amphetamine-Related Disorders	42
25	Cocaine-Related Disorders	29
26	Marijuana Abuse	59
27	Opioid-Related Disorders	402
28	Substance Abuse, Intravenous	127

29	Substance Withdrawal Syndrome	155
30	Heroin Dependence	104
31	#24 OR #25 OR #26 OR #27 OR #28 OR #29 OR #30	737
53	#32 OR #33 OR #34 OR #35 OR #36 OR #37 OR #38 OR #39 OR #40 OR #41 OR #42 OR #43 OR #44 OR #45 OR #46 OR #47 OR #48 OR #49 OR #50 OR #51 OR #52	2552
54	substance abus*	604
55	substance us*	222
56	substance addict*	198
57	drug us*	841
58	drug usage	8
59	drug abus*	266
60	drug addict*	177
61	drug depend*	108
62	#54 OR #55 OR #56 OR #57 OR #58 OR #59 OR #60 OR #61	1715
63	#7 OR #23	475
64	#31 AND #63	13
65	#62 AND #63	14
66	#63 OR #64 OR #65	475
67	#63 OR #64 OR #65 – Limited to 2015-2016.	3

The original searches for this review where conducted in cooperation between the authors and the Swedish Ministry of Health and Social Affairs.

Due to cease of cooperation, the search strategies for the following databases could not be identified and documented in the appendix:

- ASSIA – Searched until July 2013
- DiVA – Searched until July 2013
- Artikelsök – Searched until July 2013.

Grey literature search strategy

The search strategy for the grey literature search was based on the search string for the electronic database search. Due to the limited search capacity on the grey literature information resources, web pages and search engines, a shortened search string was used. The search strategy on grey literature resources, trial registers and repositories was based on a simple "two-term" search. Searches were performed by searching for "12-step*" OR "twelve-step*" in search fields on web pages. If possible, combining "12-step*" OR "twelve-step*" with "drug* OR "drug abuse*" was used. On Google Scholar, advanced search was used. "12-step*" where searched in the field *All of the words*. "Drug abuse" where searched in the field *Any of the words*. Searches were performed as title and full-text searches.

9.2 CHARACTERISTICS OF INCLUDED STUDIES

All studies analyse outcomes at the individual level. In all 10 included studies participants were individually assigned to either 12-step or the comparison condition. However some studies delivered one or more of the interventions in groups:

- Three studies reported delivering 12-step intervention individually (Carroll et al., 1998; Carroll et al., 2012; Schottenfeld et al., 2011). However in Carroll et al. (2012) 12-step was delivered in addition to a treatment as usual intervention that was group based (although participants could request individual therapy). Seven studies delivered the intervention either purely in group sessions, or a combination of individual and group session (Bisset, 2002; Gossop et al., 2007; Higgins et al., 1991; Maude-Griffin et al., 1998; McKay et al., 1997; Petry et al., 2010; Wells et al., 1994)
- Two studies reported delivering the comparison(s) individually (Carroll et al., 1998; Schottenfeld et al., 2011). In seven studies one or more of the comparisons contained a group element in delivery (Bisset et al., 2002; Carroll et al., 2012; Gossop et al., 2007; Maude-Griffin et al., 1998; McKay et al., 1997; Petry 2010; Wells et al., 1994). One study did not report whether the intervention was delivered individually or in groups (Higgins et al., 1991).
- Two studies delivered the 12-step intervention and the relevant comparisons individually (Carroll et al., 1998; Schottenfeld et al., 2011)

Below we present additional details of included studies.

Bisset, 2002

Methods	Randomised controlled trial
Participants	124 men and women from the US with an average age of 43 years old. All participants used two kinds of drugs *and* methadone when entering the trial.
Intervention	Acceptance and Commitment Therapy & Intensive Twelve Step facilitation (ITSF). Both interventions were delivered individually and in groups. Duration of interventions: 16 weeks
Outcomes and time points that are considered in this review	Drug use, criminal behaviour, psychiatric symptoms, social function and employment status. All outcomes measured 8 weeks post treatment and 6 months post treatment. Retention.
Notes	

Carroll et al., 1998

Methods	Randomised controlled trial
Participants	122 men and women with average age of 31 years old from the US were randomised. All participants had a current cocaine dependence.
Intervention	Disulfiram, Cognitive Behavioural Training & Twelve Step Facilitation. All interventions delivered individually. Duration of interventions: 12 weeks
Outcomes and time points that are considered in this review	Drug use during treatment, at the end of treatment, 1 month FU, 6 month FU, and 12 month FU. Retention. ASI subcomponent scores: post, 6 month FU, 12 month FU.
Notes	Drug use ("Days of cocaine use per month"): post, 6 month FU, 12 month FU, by author correspondence. ASI subcomponent scores: post, 6 month FU, 12 month FU by author correspondence.

Carroll et al., 2012

Methods	Randomised controlled trial
Participants	112 men and women from the US were included in the study. The average age was 38.8 overall. Participants were methadone maintenance patients.
Intervention	Disulfiram, Twelve Step Facilitation and Treatment as usual combined in different ways. Twelve step facilitation (TSF) was delivered individually. Treatment as usual delivered either in a group setting (default) or individually (participant's choice)[9]. Duration of intervention: 12 weeks
Outcomes and time points that are considered in this review	Drug use during treatment, end of treatment (post), 6 month FU, 12 months FU. Retention.
Notes	Drug use: during, post, 6 month FU, 12 month FU, by author correspondence.

[9] Source: author correspondence 7th October, 2014.

Gossop et al., 2007

Methods	Quasi-experimental study
Participants	124 men and women from the UK with an over-representation of men. The average age of the participants was 29.7 years old. All participants had a drug dependence disorder.
Intervention	Narcotics Anonymous and Alcoholics Anonymous programs. "Intervention" delivered in self-help group. Comparison is no group attendance. The programs' duration differs between 2-52 weeks
Outcomes and time points that are considered in this review	Drug use at 1 year, 2, years and 4-5 years follow-up post intake of treatment.
Notes	

Higgins et al., 1991

Methods	Quasi-experimental study
Participants	25 American men and women initiated treatment. All participants had a cocaine dependence. The mean age in the two groups was 30.5 and 29 years for the 12-step group and BT group respectively.
Intervention	Behavioural Therapy and Twelve Step Facilitation (TSF). TSF delivered in both individual and in group sessions. Behavioural Therapy delivery not reported. Duration of the programs: 12 weeks
Outcomes and time points that are considered in this review	Drug use during treatment.
Notes	Drug use. Author did not respond to our enquiries. Numerical data extracted by graphical methods.

Maude-Griffin et al., 1998

Methods	Randomised controlled trial
Participants	126 men and 2 women from the US with current cocaine abuse participated in the trial. Age not reported, but mean length of history of cocaine use in the sample was 19 years.
Intervention	Cognitive-Behavioural Coping Skills Training (CBT) and Twelve Step Facilitation (TSF). Both interventions delivered in individual and group sessions. Duration of the programs: 12 weeks
Outcomes and time points that are considered in this review	Drug use. Time points available: during, post, 14 weeks FU. Retention.
Notes	Corresponding author replied that data had been discarded. Numerical data extracted based on aggregate follow-up rates, assuming follow-up rate independent of treatment.

McKay et al., 1997

Methods	Randomised controlled trial
Participants	98 (132 in McKay 1999) men with cocaine dependence. The average age of the participants was 40 years old.
Intervention	Relapse Prevention and 12-step standard group counselling. 12-step delivered in groups. Relapse Prevention delivered in both individual and group sessions. The intervention duration for RP is unclear while the 12-step program varied in length with a maximum of 23 months counselling.
Outcomes and time points that are considered in this review	Drug use during treatment: 1-6 months (during), follow-up 7-12 months (post), 13-18 months (6 months FU), and 24 months after beginning of treatment (12 months FU). Retention. Psychiatric symptoms, employment status and family functioning at baseline, 1-6 month(during treatment) and at follow-up after 7-12 month (post), 13-18 month (6 months FU) and 24 month after treatment start (12 months FU).
Notes	Author did not respond to our enquiries. Number of observation assumed based on overall follow-up rates (assumed independent of measure and treatment).

Petry et al., 2010

Methods	Randomised controlled trial
Participants	170 HIV positive men and women with a cocaine or opioid abuse or dependence were randomised in the study. The mean age of the sample was 43.1 years and 42.6 for the 12-step group and the CM group respectively. The study was conducted in the US.
Intervention	Contingency Management and Twelve Step Facilitation (TSF). Both interventions delivered in groups. Duration of the programs: 24 weeks.
Outcomes and time points that are considered in this review	Drug use. During treatment and at 6 months follow-up. Retention.
Notes	Author did not respond to our enquiries. Drug use. Measurements taken during, post, 3 month FU, 6 months FU.

Schottenfeld et al., 2011

Methods	Randomised controlled trial
Participants	145 women from the US with a mean age of 31.1 years were randomised in the study. All the included women had a cocaine dependence and were either pregnant or had custody of a young child.
Intervention	Community reinforcement approach, Contingency management, Non-contingent- yoked voucher control and Twelve Step Facilitation (TSF). Individual delivery. Duration of the programs: 24 weeks
Outcomes and time points that are considered in this review	Drug use during treatment (week 1-24) Retention.
Notes	Author did not respond to our enquiries. Drug use. During (extracted from Fig. 3, p52), post (could not extract), 3 months FU (extracted but not used in synthesis), 6 months (could not extract) Retention. Extracted from Fig. 1, p50.

Wells et al., 1994

Methods	Quasi randomised controlled trial
Participants	110 men and women from the US with substance dependence. The median age of the sample was 28 years.
Intervention	Relapse Prevention and Twelve Step Facilitation (TSF). Both interventions delivered in groups. Duration of the programs: 24 weeks
Outcomes and time points that are considered in this review	Drug use at treatment end (post) and 6 months FU. Retention.
Notes	

9.3 CHARACTERISTICS OF EXCLUDED STUDIES

Aase , Jason, Ferrari, Li, & Scott, (2014)

Reason for Exclusion	Participants mainly alcohol abuse.

Bergman, Hoeppner, Nelson, Slaymaker, & Kelly, 2015

Reason for Exclusion	Participants. 75% cannabis users and 75% alcohol. Not reported separately.

Bogenschutz, Rice, Tonigan, Vogel, Nowinski, Hume, & Arenella, 2014

Reason for Exclusion	Participants were alcohol dependents.

Bowen, Witkiewitz, Clifasefi, Grow, Chawla, Hsu, ... , & Larimer, 2014

Reason for Exclusion	Participants. 10-15% had dependence on alcohol only.

Brooks & Penn, 2003

Reason for Exclusion	Participants. 45% of study participants have alcohol dependence disorder only. No subgroup analysis available.

Brennan, 1998

Reason for Exclusion	Participants. 52% of the participants in the treatment group and 72% of the participants in the control group had alcohol dependence disorder only. No subgroup analysis available.

Brown, Seraganian, Tremblay, & Annis, 2002 a+b

Reason for Exclusion	28.6% of the participants were alcohol dependents.

Chi, Weisner, Grella, Hser, Moore, & Mertens, 2014

Reason for Exclusion	15% alcohol dependents.

Donovan, Daley, Brigham, Hodgkins, Perl, Garrett, ... , & Kelly, 2013

Reason for Exclusion	No relevant comparison. Both interventions contain 12-step components.

Doyle, & Donovan, 2014

Reason for Exclusion	No relevant comparison. Both interventions contain 12-step components.

Fiorentine, 1999

Reason for Exclusion	At least 59 % are drug users; some participants alcohol only dependents.

Henggeler, Pickrel, & Brondino, 1999; Henggeler, Clingempeel, Brondino, & Pickrel, 2002

Reason for Exclusion	At least 68% of the participants used drugs. A majority of the participants abused or were dependent on alcohol.

Lydecker, Tate, Cummins, McQuaid, Granholm, & Brown, 2010

Reason for Exclusion	66% of the participants were only using alcohol.

Majer, Jason, Aase, Droege, & Ferrari, 2013

Reason for Exclusion	Participants. Not clear all participants were primarily drug users.

Manning, Best, Faulkner, Titherington, Morinan, Keaney, ... , & Strang, 2012

Reason for Exclusion	Participants. 57% alcohol as primary drug of use.

McKay, Lynch K G, Shepard, Ratichek, Morrison, Koppenhaver, & Pettinati, 2004

Reason for Exclusion	74.6% were cocaine dependents, the rest were only addicted to alcohol.

McKay, Lynch, Shepard, & Pettinati, 2005

Reason for Exclusion	Participants were either alcohol *or* drug addicted, outcome not specified for each addiction.

McKay, Lynch, Shepard, Morgenstern, Forman, & Pettinati, 2005

Reason for Exclusion	Participants with alcohol and/or cocaine dependence

McKay, Pettinati, Morrison, Feeley, Mulvaney, & Gallop, 2002

Reason for Exclusion	Participants were either alcohol *or* drug addicted, outcome not specified for each addiction.

Morgan-Lopez, Saavedra, Hien, Campbell, Wu, & Ruglass, 2013

Reason for Exclusion	Participants had either drug or alcohol dependence. Not reported separately.

Rosenblum, Matusow, Fong, Vogel, Uttaro, Moore, & Magura, 2014

Reason for Exclusion	Unclear whether all participants were drug dependent.

Wells, Donovan, Daley, Doyle, Brigham, Garrett, ... , & Walker, 2014

Reason for Exclusion	No relevant comparison. Both interventions contain 12-step components.

Worley, Tate, & Brown, 2012; Worley, Tate, McQuaid, Granholm, & Brown, 2013.

Reason for Exclusion	Participants were either diagnosed with alcohol, stimulant or marijuana dependence

9.4 ASSESSMENT OF RISK OF BIAS IN INCLUDED STUDIES

9.4.1 Risk of bias table

Item	**Judgement**[a]	**Description** (quote from paper, or describe key information)
1. Sequence generation		
2. Allocation concealment		
3. Confounding[b,c]		
4. Blinding?[b]		
5. Incomplete outcome data addressed?[b]		
6. Free of selective reporting?[b]		
7. Free of other bias?		
8. A priori protocol?[d]		
9. A priori analysis plan?[e]		

[a] Some items on low/high risk/unclear scale (double-line border), some on 5 point scale/unclear (single line border), some on yes/no/unclear scale (dashed border). For all items, record "unclear" if inadequate reporting prevents a judgement being made.

[b] For each outcome in the study.

[c] This item is only used for QESs. It is based on list of confounders considered important at the outset and defined in the protocol for the review (*assessment against worksheet*).

[d] Did the researchers write a protocol defining the study population, intervention and comparator, primary and other outcomes, data collection methods, etc. in advance of starting the study?

[e] Did the researchers have an analysis plan defining the primary and other outcomes, statistical methods, subgroup analyses, etc. in advance of starting the study?

9.4.2 Risk of bias tool

9.4.2.1 Studies for which RoB tool is intended

The risk of bias model is developed by Prof. Barnaby Reeves in association with the Cochrane Non-Randomised Studies Methods Group.[10] This model, an extension of the Cochrane Collaboration's risk of bias tool, covers both risk of bias in randomised controlled trials (RCTs and QRCTs), but also risk of bias in non-randomised studies (QES).

The point of departure for the risk of bias model is the Cochrane Handbook for Systematic Reviews of interventions (Higgins & Green, 2008). The existing Cochrane risk of bias tool needs elaboration when assessing non-randomised studies because, for non-randomised studies, particular attention should be paid to selection bias / risk of confounding. Additional item on confounding is used only for non-randomised studies (QESs) and is not used for randomised controlled trials (RCTs and QRCTs).

9.4.2.2 Assessment of risk of bias

Issues when using modified RoB tool to assess included non-randomised studies:

- Use existing principle: score judgement and provide information (preferably direct quote) to support judgement
- QESs.
- 5-point scale for some items (distinguish "unclear" from intermediate risk of bias).
- Keep in mind the general philosophy – assessment is not about whether researchers could have done better but about risk of bias; the assessment tool must be used in a standard way whatever the difficulty / circumstances of investigating the research question of interest and whatever the study design used.
- Anchors: "1/No/low risk" of bias should correspond to a high quality RCT. "5/high risk" of bias should correspond to a risk of bias that means the findings should not be considered (too risky, too much bias, more likely to mislead than inform)

1. Sequence generation
   - Low/high/unclear RoB item
   - Always high RoB (not random) for a non-randomised study
   - Might argue that this item redundant for QES since always high – but important to include in RoB table ('level playing field' argument)

2. Allocation concealment
   - Low/high/unclear RoB item
   - Potentially low RoB for a non-randomised study, e.g. quasi-randomised (so high RoB to sequence generation) but concealed (reviewer judges that the people making decisions about including participants didn't know how allocation was being done, e.g. odd/even date of birth/hospital number)

[10] This risk of bias model was introduced by Prof. Reeves at a workshop on risk of bias in non-randomized studies at SFI Campbell, February 2011. The model is a further development of work carried out in the Cochrane Non-Randomised Studies Method Group (NRSMG).

3. RoB from confounding (additional item for QES; assess for each outcome)
   - Assumes a pre-specified list of potential confounders defined in the protocol
   - Low(1) / 2 / 3 / 4 / high(5) / unclear RoB item
   - Judgement needs to factor in:
     - proportion of confounders (from pre-specified list) that were considered
     - whether most important confounders (from pre-specified list) were considered
     - resolution/precision with which confounders were measured
     - extent of imbalance between groups at baseline
     - care with which adjustment was done (typically a judgement about the statistical modeling carried out by authors)
   - Low RoB requires that all important confounders are balanced at baseline (not primarily/not only a statistical judgement) OR measured 'well' and 'carefully' controlled for in the analysis.

Assess against pre-specified worksheet. Reviewers will make a RoB judgement about each factor first and then 'eyeball' these for the judgement RoB table.

4. RoB from lack of blinding (assess for each outcome, as per existing RoB tool)
   - Low(1) / 2 / 3 / 4 / high(5) / unclear RoB item
   - Judgement needs to factor in:
     - nature of outcome (subjective / objective; source of information)
     - who was / was not blinded and the risk that those who were not blinded could introduce performance or detection bias
     - see Ch.8

5. RoB from incomplete outcome data (assess for each outcome, as per existing RoB tool)
   - Low(1) / 2 / 3 / 4 / high(5) / unclear RoB item
   - Judgement needs to factor in:
     - reasons for missing data
     - whether amount of missing data balanced across groups, with similar reasons
     - see Ch.8

6. RoB from selective reporting (assess for each outcome, NB different to existing Ch.8 recommendation)
   - Low(1) / 2 / 3 / 4 / high(5) /unclear RoB item
   - Judgement needs to factor in:
     - existing RoB guidance on selective outcome reporting
     - see Ch.8
     - also, extent to which analyses (and potentially other choices) could have been manipulated to bias the findings reported, e.g. choice of method of model fitting, potential confounders considered / included
     - look for evidence that there was a protocol in advance of doing any analysis / obtaining the data (difficult unless explicitly reported); QES very different from RCTs. RCTs must have a protocol in advance of starting to recruit (for REC/IRB/other regulatory

approval); QES need not (especially older studies)

- Hence, separate yes/no items asking reviewers whether they think the researchers had a pre-specified protocol and analysis plan.

Confounding worksheet

Assessment of how researchers dealt with confounding	
Method for *identifying* relevant confounders described by researchers: yes no If yes, describe the method used:	☐ ☐
Relevant confounders described: yes no List confounders described on next page	☐ ☐
Method used for controlling for confounding At design stage (e.g. matching, regression discontinuity, instrument variable): .... At analysis stage (e.g. stratification, multivariate regression, difference-indifference): .... Describe confounders controlled for below	

Confounders described by researchers

Tick (yes[0]/no[1] judgement) if confounder considered by the researchers [Cons'd?]

Score (1[good precision] to 5[poor precision]) precision with which confounder measured

Score (1[balanced] to 5[major imbalance]) imbalance between groups

Score (1[very careful] to 5[not at all careful]) care with which adjustment for confounder was carried out

Confounder	Considered	Precision	Imbalance	Adjustment
Gender	☐	☐	☐	☐
Age	☐	☐	☐	☐
Socio-economic status	☐	☐	☐	☐
Mental problems	☐	☐	☐	☐
History of drug misuse	☐	☐	☐	☐
Unobservables[11]	☐	Irrelevant	☐	☐
Other:	☐	☐	☐	☐
Other:	☐	☐	☐	☐

[11] See user guide for unobservables

9.4.3 User guide for unobservables

Selection bias is understood as systematic baseline differences between groups and can therefore compromise comparability between groups. Baseline differences can be observable (e.g. age and gender) and unobservable (to the researcher; e.g. motivation and 'ability'). There is no single non-randomised study design that always solves the selection problem. Different designs solve the selection problem under different assumptions and require different types of data. Especially how different designs deal with selection on unobservables varies. The "right" method depends on the model generating participation, i.e. assumptions about the nature of the process by which participants are selected into a program.

As there is no universal correct way to construct counterfactuals we will assess the extent to which the identifying assumptions (the assumption that makes it possible to identify the counterfactual) are explained and discussed (preferably the authors should make an effort to justify their choice of method). We will look for evidence that authors using e.g. (this is NOT a complete list):

Natural experiments:
Discuss whether they face a truly random allocation of participants and that there is no change of behaviour in anticipation of e.g. policy rules.

Instrument variable (IV):
Explain and discuss the assumption that the instrument variable does not affect outcomes other than through their effect on participation.

Matching (including propensity scores):
Explain and discuss the assumption that there is no selection on unobservables, only selection on observables.

(Multivariate) Regression:
Explain and discuss the assumption that there is no selection on unobservables, only selection on observables. Further discuss the extent to which they compare comparable people.

Regression discontinuity (RD):
Explain and discuss the assumption that there is a (strict!) RD treatment rule. It must not be changeable by the agent in an effort to obtain or avoid treatment. Continuity in the expected impact at the discontinuity is required.

Difference-in-difference (Treatment-control-before-after):
Explain and discuss the assumption that outcomes of participants and nonparticipants evolve over time in the same way.

9.5 RISK OF BIAS

Judgement score ranging from 1-4, where 1 indicates low risk of bias and 4 indicates high risk of bias. The judgement 5 indicates unacceptable high risk of bias, and the study will be excluded from the numeric analysis because of it.

Bisset 2002:

Entry		Judgement	Description
Adequate sequence generation?		Unclear	Quote: "Subjects were randomly assigned to a condition. The following variables were monitored in assigning subjects to conditions: Methadone dose, psychiatric severity, substance use severity, and methadone clinic. Subjects were then randomly assigned to one of the two therapists in each treatment condition of the study" Note: The description of the randomisation varies substantially between the two study reports. No information on how randomisation was carried out is reported.
Allocation concealment?		Unclear	Not reported who did allocation or which precautions were taken to avoid that allocation.
Blinding?	Primary outcome	3	Assessors blind to treatment. Therapists and participants were not.
	Secondary outcomes	4	
	Retention	2	
Incomplete outcome data addressed?	Primary outcome	2	Missing some intake, post intake and follow-up data.
	Secondary outcomes	3	
	Retention	1	
Selective reporting?	All outcomes	1	Data reported in table 1.
Other bias?		3	Interaction among individuals randomised to the same group therapy is expected.

		Effective sample size is smaller than reported, due to participants individually randomised to group therapy.
Followed a priori protocol?	Unclear	No registered protocol
Followed a priori analysis plan?	Unclear	No published protocol

Carroll et al., 1998:

Entry		Judgement	Description
Adequate sequence generation?		Unclear	Subjects were randomised. No further details given.
Allocation concealment?		Unclear	Not clear whether subjects, therapist, or analysists could manipulate assignment.
Blinding?	Primary outcomes	2	Neither participants nor therapists were blinded to the treatment delivered. It is not reported who measured the outcome.
	Retention	3	
Incomplete outcome data addressed?	Primary outcome	Unclear	Unclear whether measurements were taken until dropout or continued after dropout. Since mean number retained varies by condition (p718, 1998), bias is unclear. 60-70 subjects per time point followed up. Completers more likely to respond.
	Secondary outcomes	Unclear	
	Retention	Unclear	
Selective reporting?	Primary outcome	2	Administer the ASI at baseline. Data retrieved via author correspondence.
	Secondary outcomes	1	
	Retention	2	Not reported by treatment
Other bias?		1	

Followed a priori protocol?	Unclear	Not reported
Followed a priori analysis plan?	Unclear	Not reported

Carroll et al., 2012:

Entry		Judgement	Description
Adequate sequence generation?		Low risk	Computer urn algorithm from project MATCH.
Allocation concealment?		Low risk	Each new draw is a biased coin based on assignment of previous patient.
Blinding?	Primary outcome	3	Subjects blind to Disulfiram/Placebo comparison, probably not blind to tau+tsf/tau. Personnel non-blind to tau+tsf/tau comparison, unclear whether blind to Dis/Pla comparison.
	Retention	3	
Incomplete outcome data addressed?	Primary outcome	2	Complete self-report outcome data available for 98% during treatment and 93% at 12 months follow-up
	Retention	1	
Selective reporting?	Primary outcome	1	
	Retention	2	Details are not reported for estimates which are statistically insignificant
Other bias?		4	TAU is group based intervention, and 3. order bias due to urn procedure.
Followed a priori protocol?		Yes	clinicaltrials.gov id NCT00350870
Followed a priori analysis plan?		Unclear	Not stated

Gossop et al., 2007:

Entry		Judgement	Description
Adequate sequence generation?		High risk	The study is not a randomised study
Allocation concealment?		High risk	High risk of selection bias
Confounding		5	Balance not displayed, cannot judge distributional overlap. Sites are different based on descriptive piece, but not controlled for. Control for pre-intake NA/AA attendance and dependence severity. Over time hardly any change in the number subjects attending NA/AA. Pre exposure is strongly correlated with post take up (p.121), which suggest that identification of the effect of attendance comes primarily from those individual who attended and subsequently dropped out.
Blinding?	Primary outcomes	4	Not blind
Incomplete outcome data addressed?	Primary outcome	3	Eligible sample of 255. Analysis sample is 142 subjects who had data for all years. Attrition analysis suggests 113 subjects have similar pre-intake characteristics to the analysis sample.
Selective reporting?	Primary outcome	1	No sign of selective reporting
Other bias?		3	Treatment is 12-step attendance, a group based intervention. Risk of bias since participants may interact during treatment (should be cluster corrected on group identifier).
Followed a priori protocol?		Unclear	Referenced as a prospective cohort study. No approved protocol.
Followed a priori analysis plan?		Unclear	Not reported

Higgins et al., 1991:

Entry		Judgement	Description
Adequate sequence generation?		High risk	Non-random sequence generation. 13 consecutive to Behavioural Therapy, next 15 to 12-step
Allocation concealment?		High risk	The fact that recruitment to BT stops after 13 (not pre-specified) suggests that allocation was not concealed to either staff or analyst.
Blinding?	Primary outcome	4	Not blinded. Since Behavioural therapy is linked to incentive pay, it needs to be clarified who decided what. Appears that staff played a role in linking performance to pay. Retention: same concerns as for drug use, but smaller risk of bias due to the nature of the outcome.
	Retention	3	
Incomplete outcome data addressed?	Primary outcome	3	Treat missing UAs as positive screen for drug use. However they only make payment based on cocaine screen. Appears to be drug substitution in the BT treatment. Appears to be differential retention between behavioural and 12-step is a concern. 85% completed vs 42% Missing data level not reported.
	Retention	1	
Selective reporting?	Primary outcome	2	Missing data level not reported
Other bias?		3	12-step is group based.
Followed a priori protocol?		Unclear	Not stated
Followed a priori analysis plan?		Unclear	Not stated

Maude-griffin et al., 1998:

Entry		Judgement	Description
Adequate sequence generation?		Unclear	There are no details on the randomisation process.
Allocation concealment?		Unclear	No details given
Blinding?	Primary outcome	Unclear	It is not stated who did the assessment. Therapists not blinded, and were rotated and delivered therefore both types of treatments. Subjects not blind.
	Retention	Unclear	
Incomplete outcome data addressed?	Primary outcome	2	89% completed 4 week, 86% 8 week, 92% 12 week, 84% 26 weeks. No differential attrition.
	Retention	1	
Selective reporting?	Primary outcome	1	Retention is not reported by treatment condition
	Retention	2	
Other bias?		3	The participants received group based intervention. Not cluster corrected on group identifier.
Followed a priori protocol?		Unclear	Not reported
Followed a priori analysis plan?		Unclear	Not reported

McKay et al., 1997:

Entry		Judgement	Description
Adequate sequence generation?		Low risk	Participants randomised by urn randomisation
Allocation concealment?		Unclear	Insufficient details given
Blinding?	Primary outcome	4	Neither participants nor clinicians are blinded to treatment. Baseline

	Secondary outcomes	3	and follow-up assessors had had been informed of treatment condition. ASI interview used to secondary outcomes were non blinded
	Retention	3	
Incomplete outcome data addressed?	Primary outcome	2	Follow-up rate not specified separately by outcome. 100% baseline, 98% 6 months, 92% 24 months. Urine toxicology results were obtained from 93 patients (95%) at 3 months and 91 patients (93%) at 6 months.
	Retention	1	
Selective reporting?	All outcomes	1	Retention is not reported by treatment condition
Other bias?		4	Both treatments contain group session element, which is not corrected for.
Followed a priori protocol?		Unclear	Not reported.
Followed a priori analysis plan?		Unclear	Not stated. ASI measures only done post treatment, and at baseline, so likely not in original protocol.

Petry et al., 2010:

Entry	Judgement	Description
Adequate sequence generation?	Low risk	Quote: "Randomly assigned to of two treatments and scheduled to attend first group session within a 1-week period. Used a computerised urn procedure that balanced treatment groups on race, baseline urine analysis, and HIV medication". Appears to be a minimisation algorithm.
Allocation concealment?	Low risk	Computer algorithm likely a minimisation algorithm.

Blinding?	Primary outcome	4	Outcome collected by research assistant, who were not was blinded.
	Retention	3	According to retention, nothing is stated but likely taken by therapists who were not blind to treatment.
Incomplete outcome data addressed?	Primary outcome	4	Primary: Large discrepancy between figure 1 and p. 92.
	Retention	1	
Selective reporting?	Primary outcome	3	Outcomes not reported for 6 months post follow-up and 9 months follow-up of screens.
	Retention	1	Do not report numbers of participants according to retentions rate.
Other bias?		4	Not corrected for group based intervention. Screens were collected weekly, since cocaine metabolises quickly there is the possibility to game the screens. 4 year recruitment period, possibility of time trends in treatment, and or participant characteristics.
Followed a priori protocol?		Unclear	Participants signed consent forms, but it is not stated whether protocol was approved in advance
Followed a priori analysis plan?		Unclear	Unclear

Schottenfeld et al., 2011:

Entry	Judgement	Description
Adequate sequence generation?	Low risk	Urn randomisation was utilised.
Allocation concealment?	Low risk	Computer algorithm likely a minimisation algorithm. Not stated

			whether allocation was done by third part.
Blinding?	Primary outcome	4	Participants were not blinded to the treatment, neither were the therapists. According to retention, double non-blind. In particular risk that therapists favouring a treatment may work differentially harder at retaining patients.
	Retention	3	
Incomplete outcome data addressed?	Primary outcome	3	Assessment completion rates were 86% (n=125) for the midpoint of treatment (3-month), 72% (n=105) for the end of treatment (6-month), and 66% (n=96) for one of the post-treament follow-up assessments. 49% (n=71) at month 9 and 48% (n=69) at month 12.
	Retention	1	
Selective reporting?	All outcomes	1	Reports on all 3 measures related to cocaine use.
Other bias?		2	Women assigned to 12-step reported significantly more days of use than those assigned to CRA. Suggesting randomisation did not fully work on an important (confounding) variable. Urn randomisation leads to risk of 3rd order bias. Makes it likely that those assigning treatment could predict with some precision the following allocation (biased urn).
Followed a priori protocol?		Yes	Protocol approved by Yale Medical School's Human Investigation Committee. Trial registered at Clinicaltrials.gov as NCT00914381
Followed a priori analysis plan?		Unclear	Not reported. Reviewer checked protocol, but it contains no analysis plan.

Wells et al., 1994:

Entry		Judgement	Description
Adequate sequence generation?		High risk	Alternation, and in addition 48 assigned to RP and 62 to TS. Authors note this is done because TS had higher attrition.
Allocation concealment?		High risk	Since alternation was the allocation method it is likely that both subjects and researchers could predict the next assignment.
Blinding?	Primary outcome	4	Participants and therapists non-blind to treatment. Not reported who conducted interviews. Risk of differential effect from therapist non-blinding on retention.
	Retention	3	
Incomplete outcome data addressed?	Primary outcome	3	17 subjects lost to follow-up. Obtained complete data on cocaine use at 3 measurement points for 92 subjects (42 Rel, 50 in TS). Only those 92 are used for analysis. 110 assigned so one is missing. Attrition 6 from RP, and 12 from TS. Retention is not listed as outcome, but presented in results section as being contrary to expectation, and there only reported as mean number of sessions (number of observations not reported.
	Secondary	3	
	Retention	Unclear	
Selective reporting?	Primary outcome	2	Only partial reporting on the complete abstinence outcome. Table 2 only reports on days of drug use. Urine screens not reported either (although listed as validation tool only).
	Retention	Unclear	
Other bias?		4	No therapist crossover increases risk of therapist effects driving

		results. Both interventions are group based, but no correction made in analysis. Recruitment procedure itself appears to have a long temporal aspect increasing the risk that changes in the participant pool or treatment fidelity over time may confound results.
Followed a priori protocol?	Unclear	Not stated
Followed a priori analysis plan?	Unclear	Not stated

10 Figures

10.1 NARCOTICS ANONYMOUS

The 12 Steps of Narcotics Anonymous (Narcotics Anonymous, 2008)

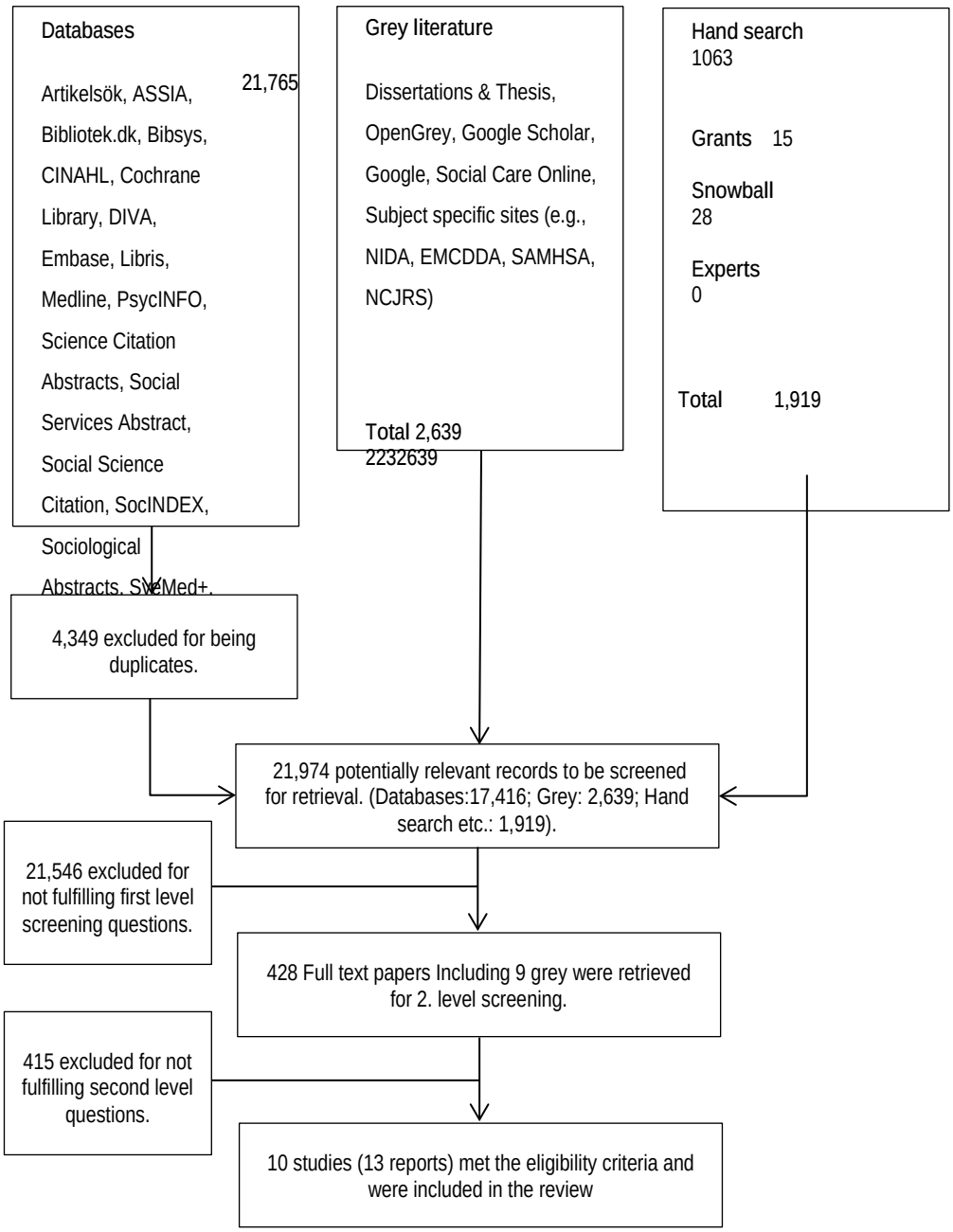
- We admitted that we were powerless over our addiction, that our lives had become unmanageable.
- We came to believe that a power greater than ourselves could restore us to sanity.
- We made a decision to turn our will and our lives over to the care of God as we understand Him.
- We made a searching and fearless moral inventory of ourselves.
- We admitted to God, to ourselves, and to another human being the exact nature of our wrongs.
- We were entirely ready to have God remove all these defects of character.
- We humbly asked him to remove our shortcomings.
- We made a list of all persons we had harmed, and became willing to make amends to them all.
- We made direct amends to such people wherever possible, except when to do so would injure them or others.
- We continued to take personal inventory when we were wrong and promptly admitted it.
- We sought through prayer and meditation to improve our conscious contact with God as we understood Him, praying only for knowledge of His will for us and the power to carry that out.
- Having had a spiritual awakening as the result of these steps, we tried to carry this message to addicts, and to practice these principles in all our affairs.

The 12 Traditions of Narcotics Anonymous (Narcotics Anonymous, 2008)

- Our common welfare should come first; personal recovery depends on NA unity.
- For our group purpose there is but one ultimate authority – a loving God as He may express Himself in our group conscience. Our leaders are but trusted servants; they do not govern.
- The only requirement for membership is a desire to stop using.
- Each group should be autonomous except in matters affecting other groups or NA as a whole.
- Each group has but one primary purpose – to carry the message to the addict who still suffers.
- An NA group ought never endorse, finance, or lend the NA name to any related facility or outside enterprise. Lest problems of money, property, or prestige divert us from our primary purpose.
- Every NA group ought to be fully self-supporting, declining outside contributions.
- Narcotics Anonymous should remain forever nonprofessional, but our service centres may employ special workers.
- NA, as such, ought never be organised, but we may create service boards or committees directly responsible to those they serve.
- Narcotics Anonymous has no opinion on outside issues: hence the NA name ought never be drawn into public controversy.
- Our public relations policy is based on attraction rather than promotion; we need always maintain personal anonymity at the level of press, radio, and films.
- Anonymity is the spiritual foundation of all our Traditions, ever reminding us to place principles before personalities.

10.2 FLOW OF STUDIES

Figure 2 Flowchart



11 Data and analyses

11.1 DESCRIPTION OF MEASURES

Measure	Description
Primary Outcome: Drug use	
Bisset (2002)	*Measure(s):* abstinence *Instrument(s):* urine analysis screen. Urine was screened for: “[...] polydrug use, including opiates, cocaine, benzodiazepines, barbiturates, amphetamines, marijuana, alcohol, and methadone" (Bisset, 2002, p.76); “relevant portions of ASI” (Hayes et al., 2004, p 673) *Time Point(s):* baseline, 8 weeks, post treatment (16 weeks), and FU (6mths post treatment). Urine Analysis collected twice per week at group and individual session
Carroll et al. (1998)	*Measure(s):* duration of periods of abstinence from cocaine, alcohol and both substances; frequency of cocaine use (number of days per week); quantity use (grams per week). *Instrument(s):* self-reports verified by UA. In cases of discrepancy the instrument indicating cocaine use was used. Follow-up: Days of use in past 28 days based Substance Abuse Calendar. *Time Point(s):* Baseline, weekly during treat, post (12 weeks), and FU: 1, 3, 6, 12 months post.
Carroll et al. (2012)	*Measure(s):* % days of cocaine use/number of days per week cocaine use; % cocaine-positive urine samples. *Instrument(s):* Self-report (Substance Abuse Calendar); urine toxicology screen *Time Point(s):*

	Self-report: Baseline, weekly during treatment, treatment end (12 weeks), every 3 months post treatment (up to 1 year); urine: 3 times weekly during treatment, and at each FU.
Gossop et al. (2007)	*Measure(s):* substance use for the 90 previous days *Instrument(s):* Structured interview *Time Point(s):* 1, 2, and 4-5 years post intake to treatment
Higgins et al. (1991)	*Measure(s):* abstinence *Instrument(s):* Urine screen. Screened for metabolite of cocaine, and one other randomly selected specimen per week screened for other drugs of abuse *Time Point(s):* collected 4 times per week during treatment
Maude-Griffin et al. (1998)	*Measure(s):* 4 consecutive weeks of abstinence during treatment verified via urine samples; point prevalence at each assessment. *Instrument(s):* urine samples collected once each week on a random schedule. To be coded as abstinent, the participant had to report no cocaine use during the prior 30 days and produce a cocaine-free urine sample. Any discrepancy coded as not abstinent. *Time Point(s):* baseline, weeks 4, 8, 12, and 26
McKay et al. (1997, 1999)	*Measure(s):* percent days of cocaine use. *Instrument(s):* Timeline Followback; ASI-drug; urine screens for validation. *Time Point(s):* 3, 6, 12, 18, and 24 months post intake.
Petry et al. (2010)	*Measure(s):* longest consecutive number of weeks of negative samples submitted; proportion of negative samples submitted. *Instrument(s):* urine screens. *Time Point(s):* weekly during treatment screened for cocaine and opioids; 3, 6 (post treatment), 9, 12 months

Schottenfeld et al. (2011)	*Measure(s):* maximum consecutive weeks of documented cocaine abstinence; the proportion of negative urine tests; percent days using cocaine; FU: 30 days cocaine abstinence *Instrument(s):* weekly TLFB of drug use (Weekly Substance Use Inventory); ASI: drug composite score. *Time Point(s):* twice weekly during treatment; FU assessment: urine toxicology and ASI. FU at 3, 6, 9, and 12 months. (based on self-report + negative urine sample)
Wells et al. (1994)	*Measure(s):* number days of use in past 30 days (cocaine, alcohol, marijuana) *Instrument(s):* monthly random urine samples *Time Point(s):* baseline, 12 weeks (post treatment), 6 months after treatment end.

Measure	Description
Secondary outcome: Retention	
Bisset (2002)	ACT/12-step: attending at least 50% of the group and individual sessions (p675); MM: giving urine samples at least 8/16 weeks
Carroll et al. (1998)	Treatment completers
Carroll et al. (2012)	Treatment completers; days retained in treatment
Gossop et al. (2007)	Not relevant (only completers)
Higgins et al. (1991)	Treatment completers
Maude-Griffin et al. (1998)	Therapy session attendance
McKay et al. (1997, 1999)	number of continuing care sessions attended during treatment
Petry et al. (2010)	Number of sessions attended
Schottenfeld et al. (2011)	Retention in treatment
Wells et al. (1994)	Number of sessions attended

Measure	Description
Secondary outcome: Criminal behaviour, psychiatric symptoms, social functioning,	

employment status	
Bisset (2002)	ASI-legal; Social Adjustment Scale, Beck Depression Index, Symptom Checklist-90-R, ASI-psychiatric; ASI-family; ASI-employment
Carroll et al. (1998)	ASI-legal; ASI-psychiatric; ASI-family; ASI-employment
McKay et al. (1997, 1999)	ASI-psychiatric; ASI-family; ASI-employment

11.2 OUTCOME MEASURE BY STUDY AND TIME OF MEASUREMENT

Table 15: Type of outcome measure by study and time of measurement

Study	Outcome					
	Primary: Drug use		Secondary: Retention		Secondary: Other	
	Continuous	Discrete	Continuous	Discrete	Continuous	Discrete
Bisset 2002	-	**po, fu**	-	po	po, fu	-
Carroll 1998	du, po, fu	du	-	po	po, fu	-
Carroll 2012	du, po, fu	-	po	po	-	-
Higgins 1991	du	du	-	po	-	-
Maude-Griffin 1998	-	**du, po, fu**	-	-	-	-
McKay 1997	du, po, fu	-	**po**	-	po, fu	-
Petry 2010	du	**fu**	**po**	-	-	-
Schottenfeld 2011	du	**fu**	-	po	-	-
Wells 1994	po, fu	-	**po**	-	-	-

Notes: Time of measurement: du = during, po = post, fu = follow-up. Effect size transformed for synthesis (from discrete to continuous or vice versa) in bold.

11.3 STUDY EFFECT SIZES

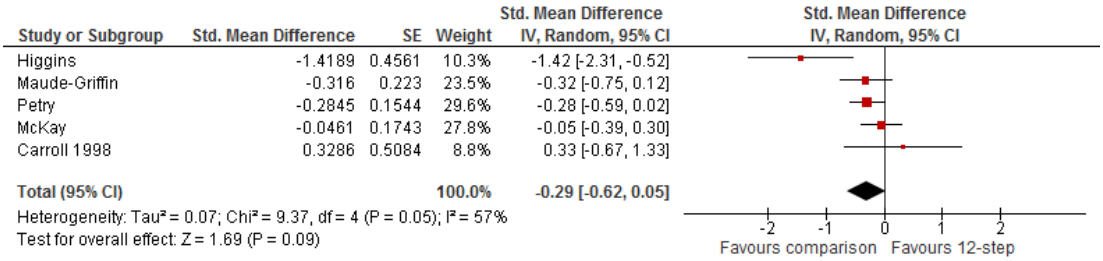
Effect sizes for each study is provided in the online appendix, see https://campbellcollaboration.org/library/12-step-programmes-illicit-drug-abuse-reduction.html.

11.4 FOREST PLOTS

11.4.1 Comparison: 12-step vs psychosocial intervention

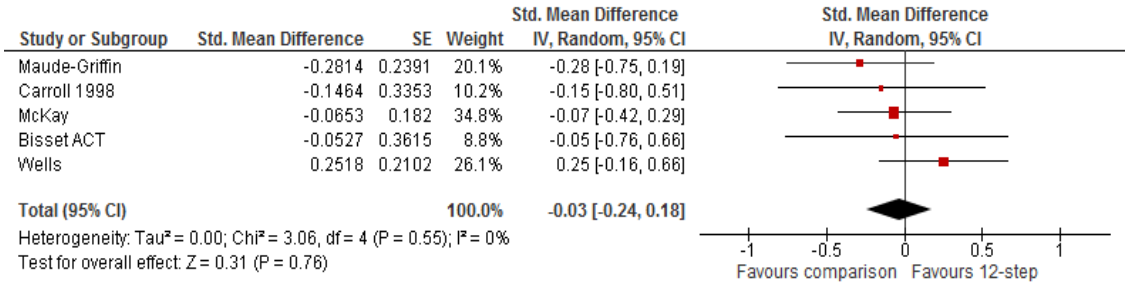
11.4.1.1 Primary outcome: Drug use

11.4.1.1.1 During treatment



A positive effect size favours 12-step.

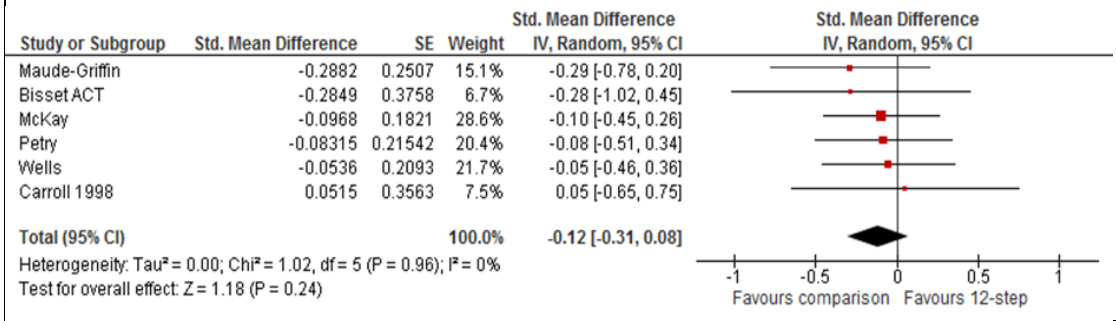
11.4.1.1.2 Post treatment



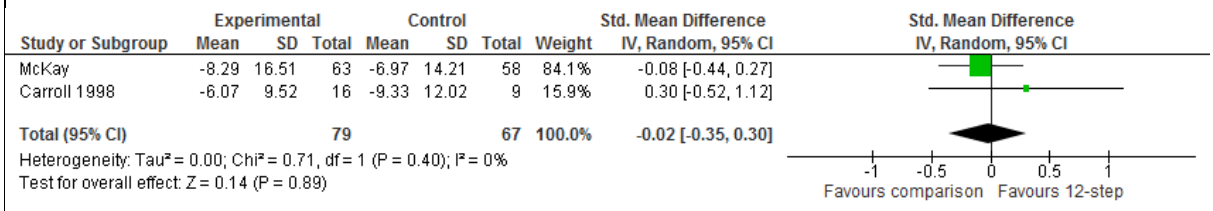
A positive effect size favours 12-step.

11.4.1.1.3 Follow-up

FU 6m



FU 12-18m

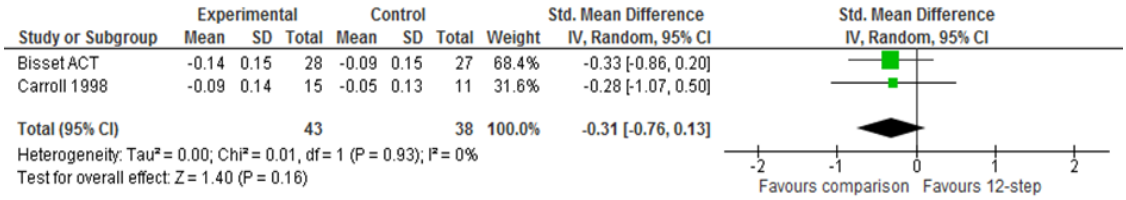


A positive effect size favours 12-step.

11.4.1.2 Secondary outcomes

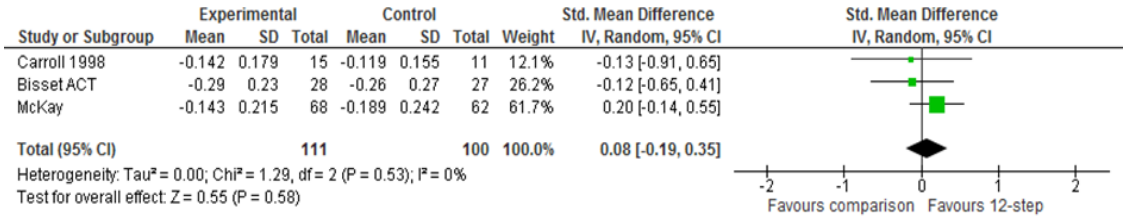
11.4.1.2.1 Post treatment

Criminal behaviour



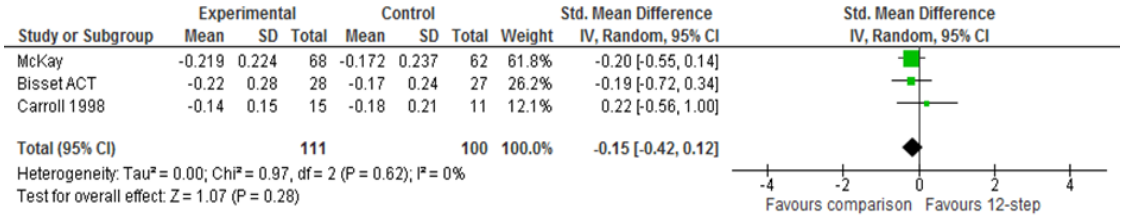
A positive effect size favours 12-step.

Psychiatric symptoms



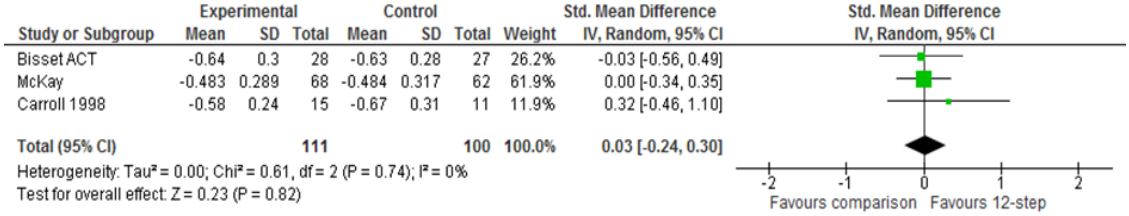
A positive effect size favours 12-step.

Social functioning



A positive effect size favours 12-step.

Employment

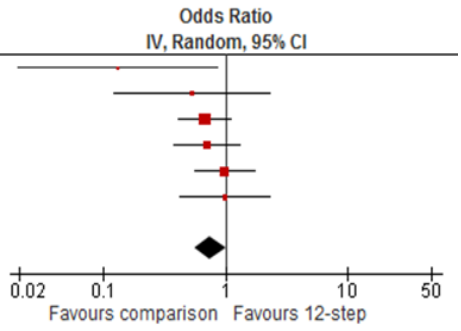


A positive effect size favours 12-step.

Retention

Study or Subgroup	log[Odds Ratio]	SE	Weight	Odds Ratio IV, Random, 95% CI
Higgins	-2.0412	0.9663	2.4%	0.13 [0.02, 0.86]
Carroll 1998	-0.6286	0.7544	3.9%	0.53 [0.12, 2.34]
Petry	-0.396	0.2527	34.1%	0.67 [0.41, 1.10]
Wells	-0.3465	0.32	21.4%	0.71 [0.38, 1.32]
McKay	-0.0165	0.2864	26.6%	0.98 [0.56, 1.72]
Bisset ACT	-0.0087	0.4333	11.7%	0.99 [0.42, 2.32]
Total (95% CI)			**100.0%**	**0.75 [0.56, 1.00]**

Heterogeneity: Tau² = 0.00; Chi² = 5.03, df = 5 (P = 0.41); I² = 1%
Test for overall effect: Z = 1.93 (P = 0.05)



An effect size > 1 favours 12-step.

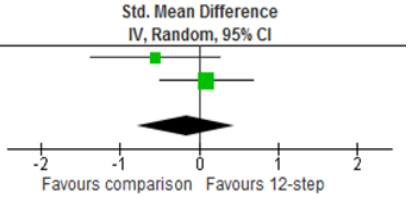
11.4.1.2.2 Follow-up

Criminal behaviour

6m FU

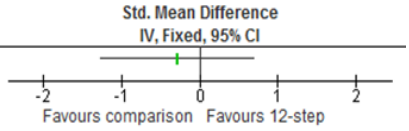
Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Carroll 1998	-0.09	0.17	17	-0.01	0.03	9	39.6%	-0.55 [-1.38, 0.27]
Bisset ACT	-0.03	0.11	26	-0.04	0.11	19	60.4%	0.09 [-0.50, 0.68]
Total (95% CI)			**43**			**28**	**100.0%**	**-0.17 [-0.78, 0.45]**

Heterogeneity: Tau² = 0.07; Chi² = 1.54, df = 1 (P = 0.21); I² = 35%
Test for overall effect: Z = 0.53 (P = 0.60)



12m FU

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Carrol 1998 12 m FU	-0.01	0.04	13	-0	0	6		-0.28 [-1.26, 0.69]



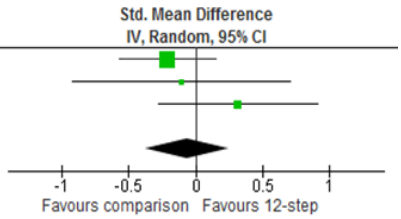
A positive effect size favours 12-step.

Psychiatric symptoms

6m FU

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
McKay	-0.162	0.226	63	-0.116	0.207	58	60.4%	-0.21 [-0.57, 0.15]
Carroll 1998	-0.21	0.19	17	-0.19	0.17	9	14.3%	-0.11 [-0.91, 0.70]
Bisset ACT	-0.26	0.26	26	-0.34	0.23	19	25.3%	0.32 [-0.28, 0.91]
Total (95% CI)			**106**			**86**	**100.0%**	**-0.06 [-0.38, 0.25]**

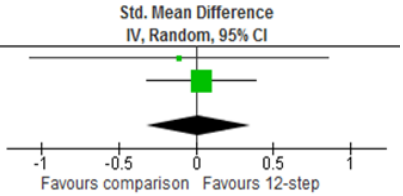
Heterogeneity: Tau² = 0.01; Chi² = 2.22, df = 2 (P = 0.33); I² = 10%
Test for overall effect: Z = 0.39 (P = 0.70)



12m FU

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Carroll 1998	-0.18	0.18	13	-0.16	0.15	6	12.0%	-0.11 [-1.08, 0.86]
McKay	-0.147	0.225	63	-0.155	0.24	58	88.0%	0.03 [-0.32, 0.39]
Total (95% CI)			**76**			**64**	**100.0%**	**0.02 [-0.32, 0.35]**

Heterogeneity: Tau² = 0.00; Chi² = 0.08, df = 1 (P = 0.78); I² = 0%
Test for overall effect: Z = 0.10 (P = 0.92)



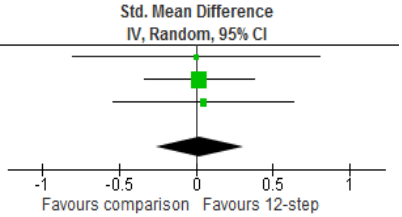
A positive effect size favours 12-step.

Social functioning

6m FU

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Carroll 1998	-0.15	0.21	17	-0.15	0.16	9	12.5%	0.00 [-0.81, 0.81]
McKay	-0.162	0.192	63	-0.166	0.21	58	64.2%	0.02 [-0.34, 0.38]
Bisset ACT	-0.14	0.21	26	-0.15	0.19	19	23.3%	0.05 [-0.54, 0.64]
Total (95% CI)			**106**			**86**	**100.0%**	**0.02 [-0.26, 0.31]**

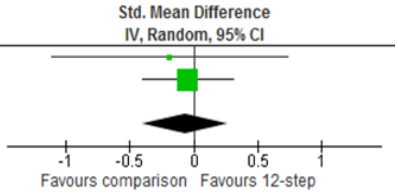
Heterogeneity: Tau² = 0.00; Chi² = 0.01, df = 2 (P = 0.99); I² = 0%
Test for overall effect: Z = 0.16 (P = 0.87)



12m FU

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Carroll 1998	-0.13	0.21	13	-0.09	0.19	7	13.0%	-0.19 [-1.11, 0.73]
McKay	-0.161	0.204	63	-0.152	0.188	58	87.0%	-0.05 [-0.40, 0.31]
Total (95% CI)			**76**			**65**	**100.0%**	**-0.06 [-0.40, 0.27]**

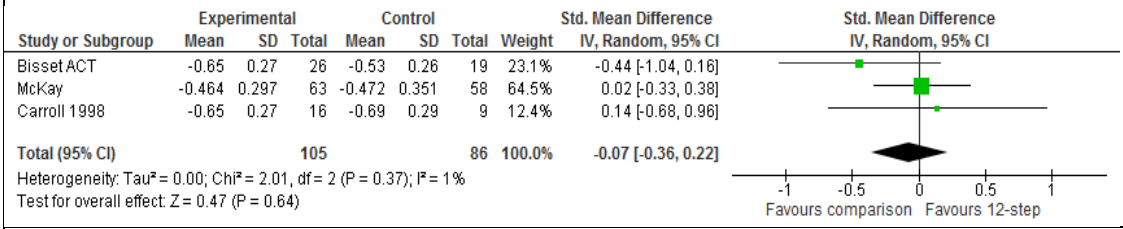
Heterogeneity: Tau² = 0.00; Chi² = 0.08, df = 1 (P = 0.78); I² = 0%
Test for overall effect: Z = 0.38 (P = 0.71)



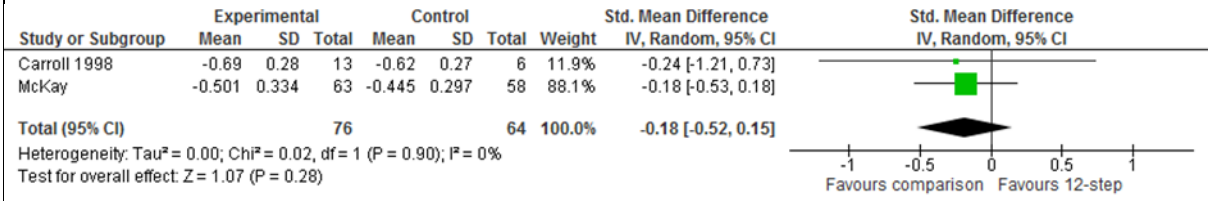
A positive effect size favours 12-step.

Employment

6m FU



12m FU

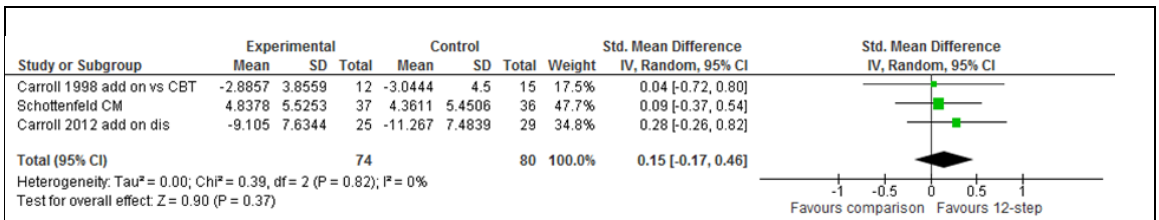


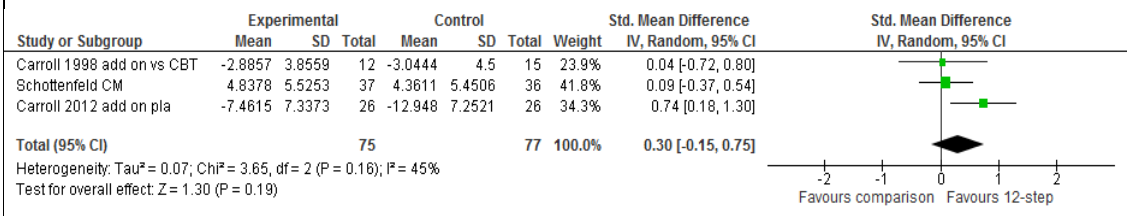
A positive effect size favours 12-step.

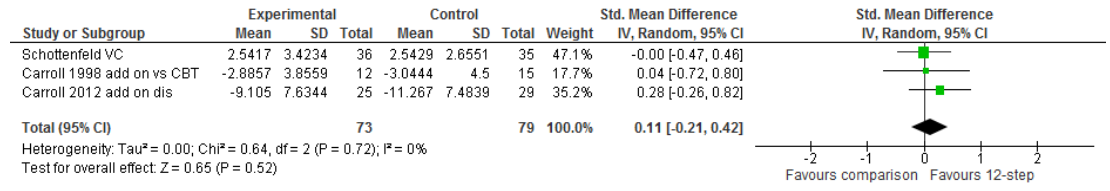
11.4.2 Comparison: 12-step + add-on vs psychosocial + add-on

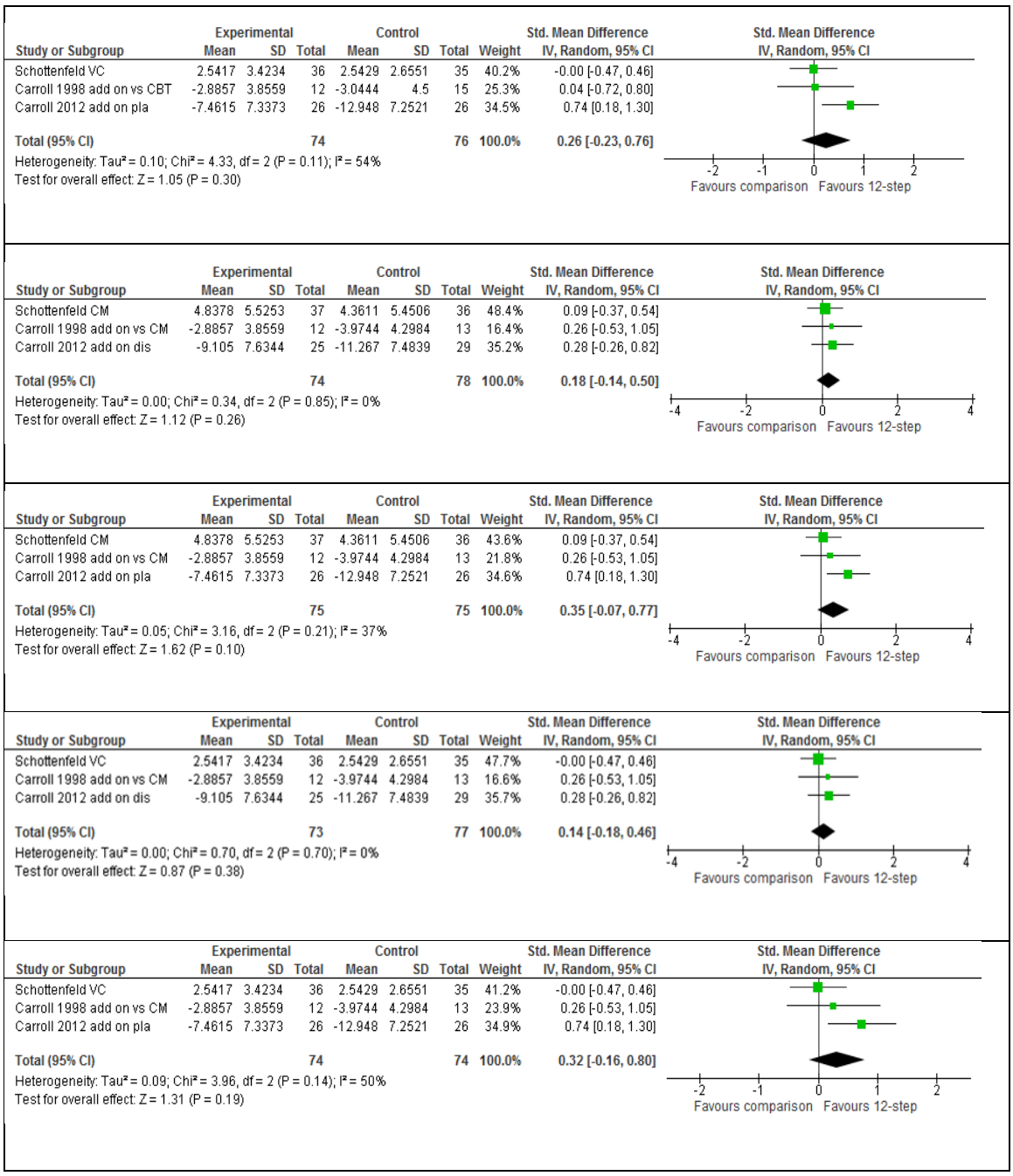
11.4.2.1 Primary outcome: Drug use

11.4.2.1.1 During treatment



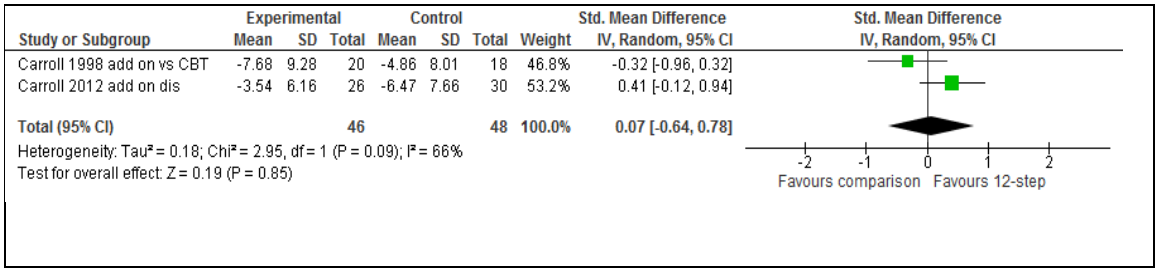


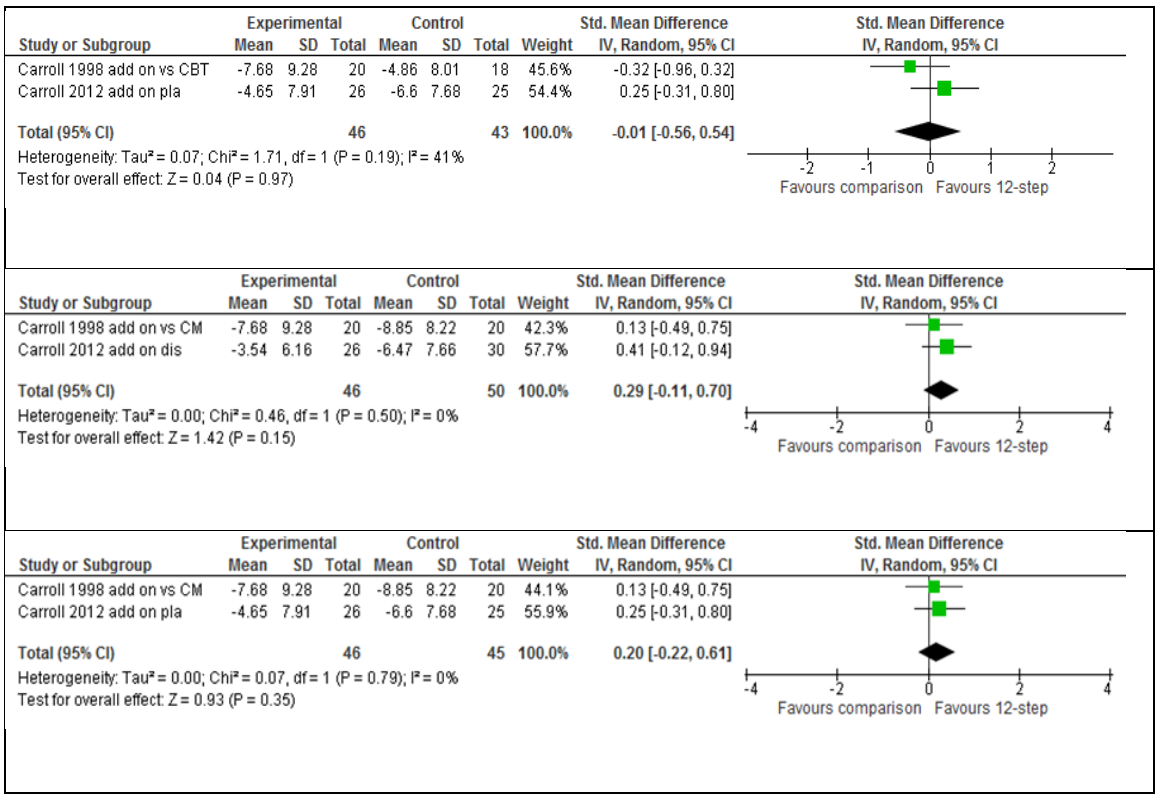




A positive effect size favours 12-step.

11.4.2.1.2 Post treatment

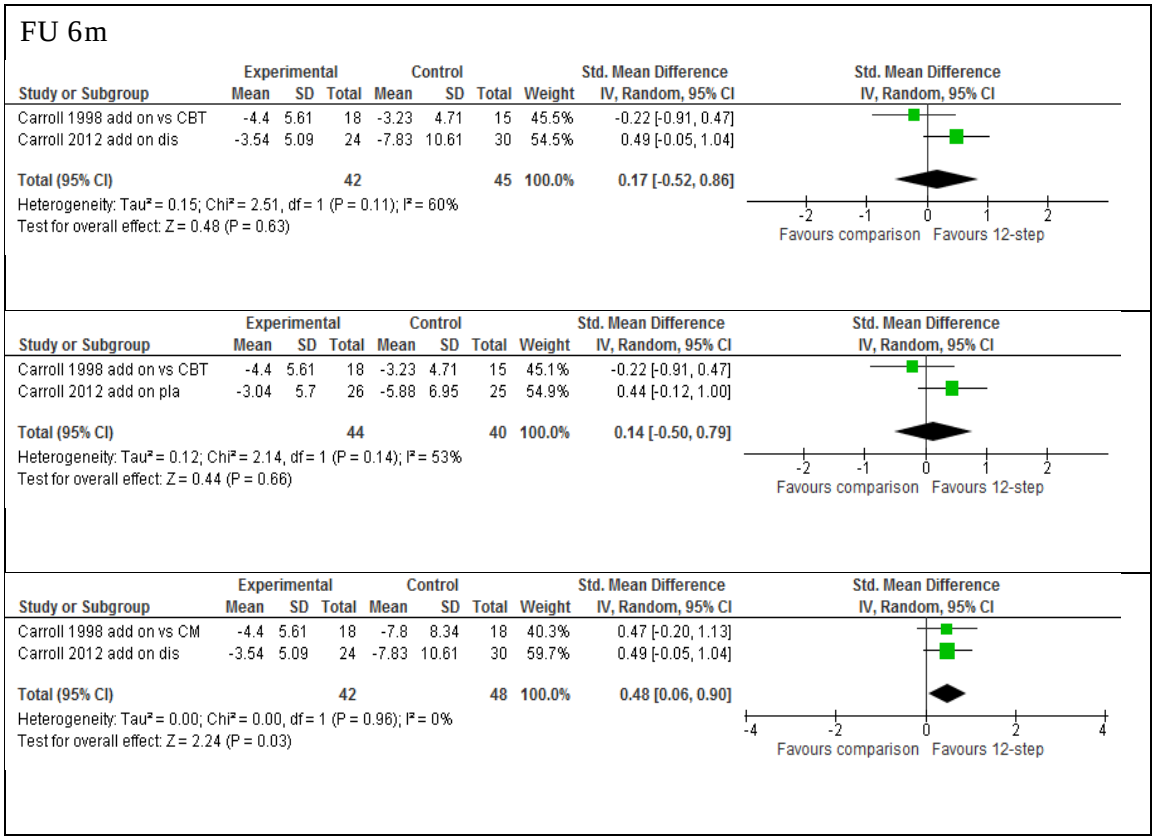


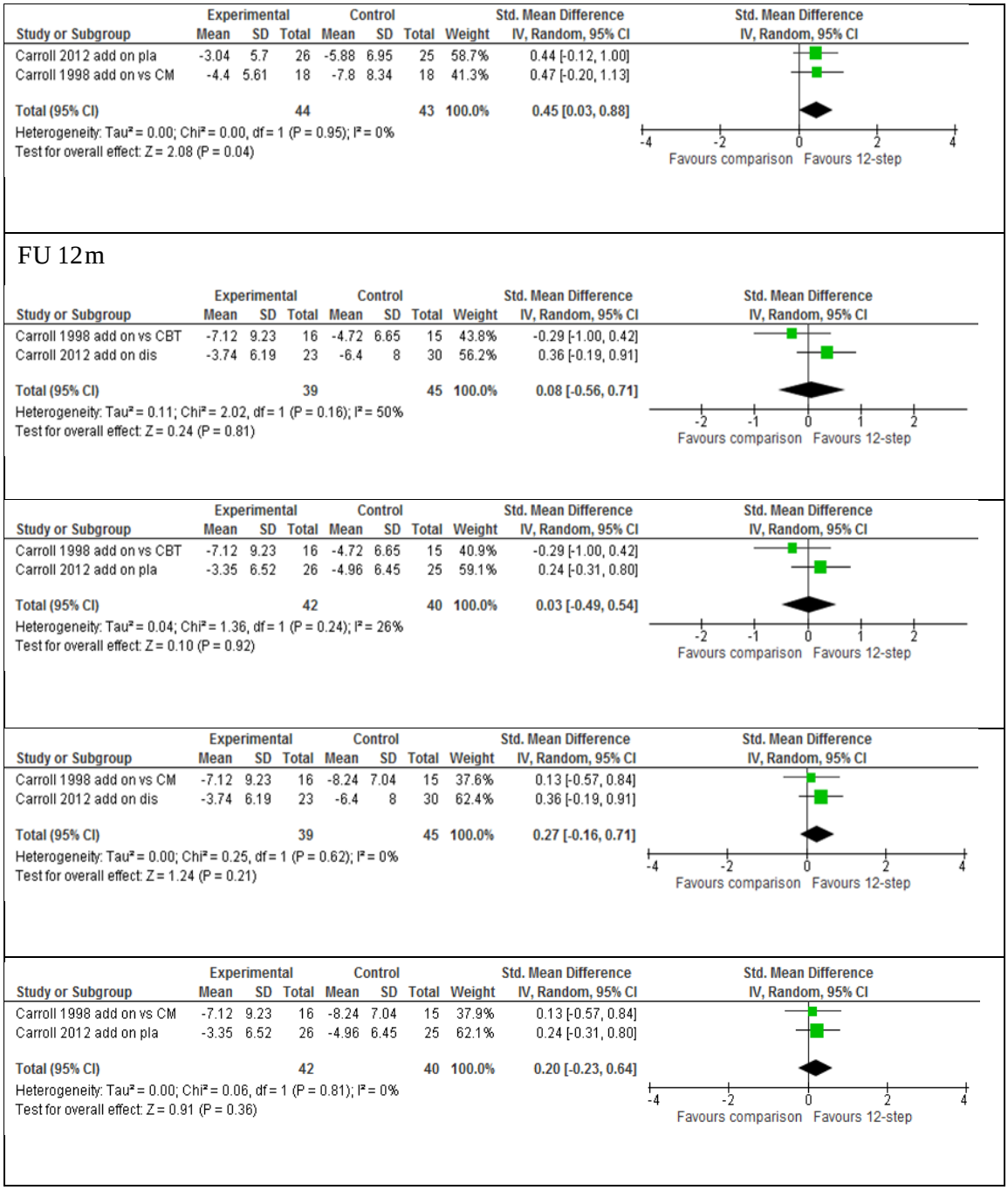


A positive effect size favours 12-step.

11.4.2.1.3 Follow-up

FU 6m



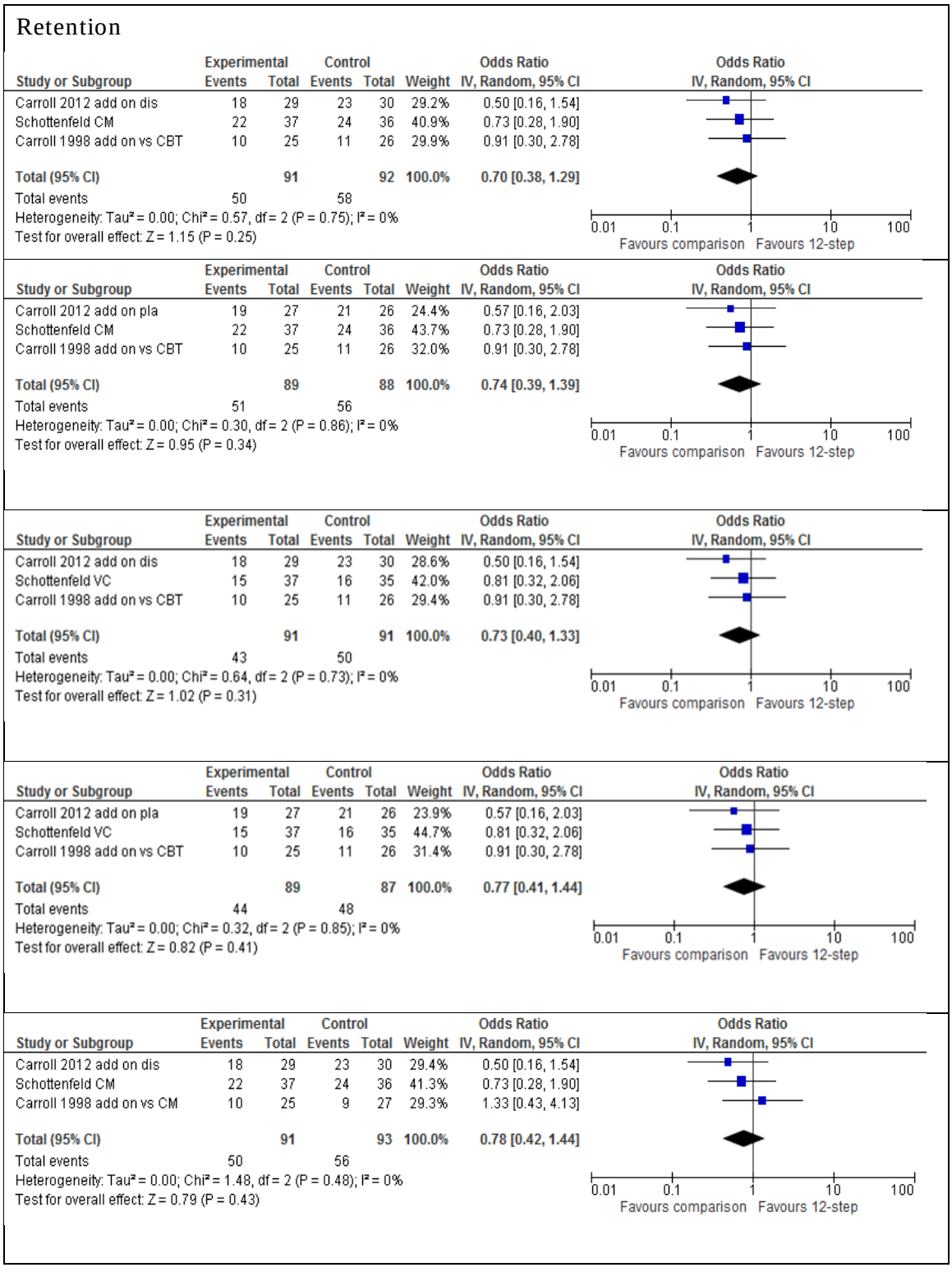


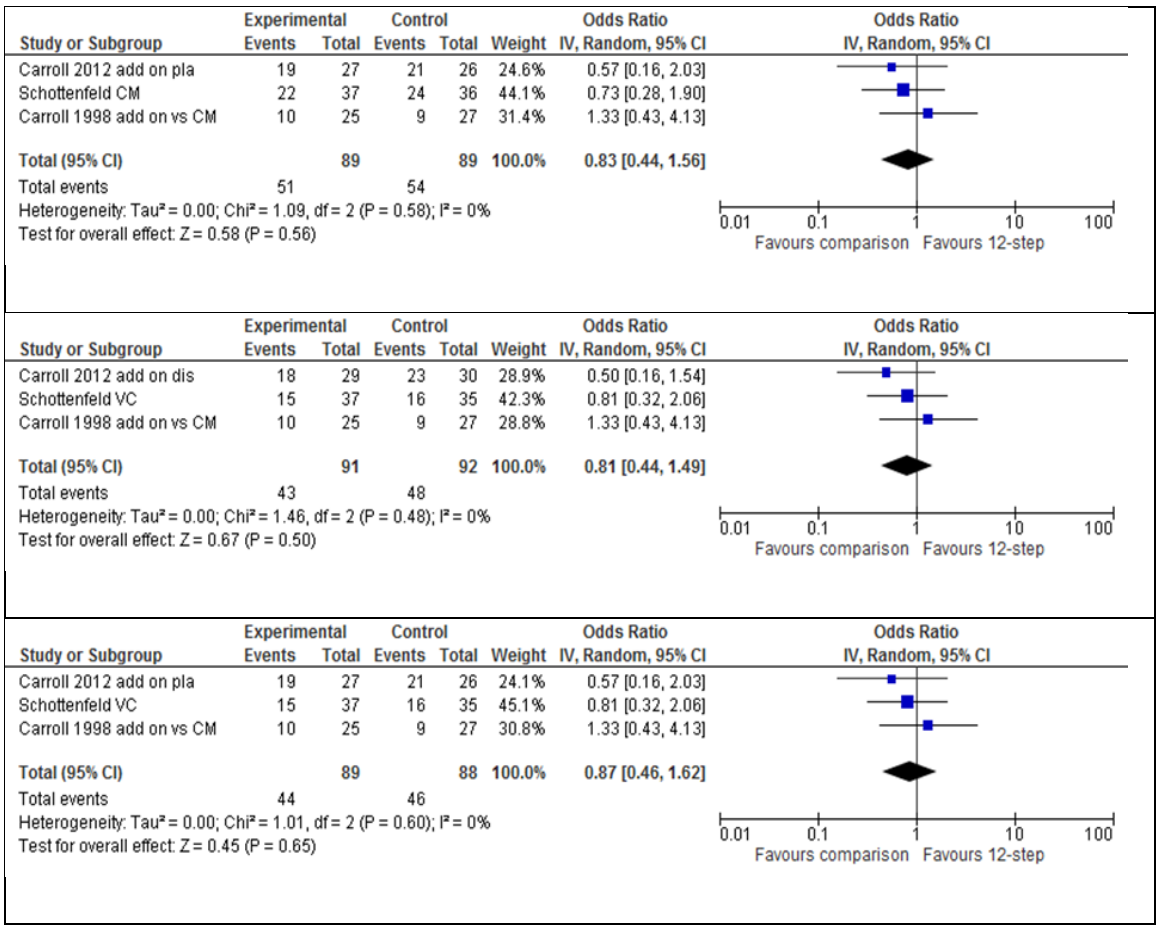
A positive effect size favours 12-step.

11.4.2.2 Secondary outcomes

11.4.2.2.1 Post treatment

Retention





An effect size > 1 favours 12-step.

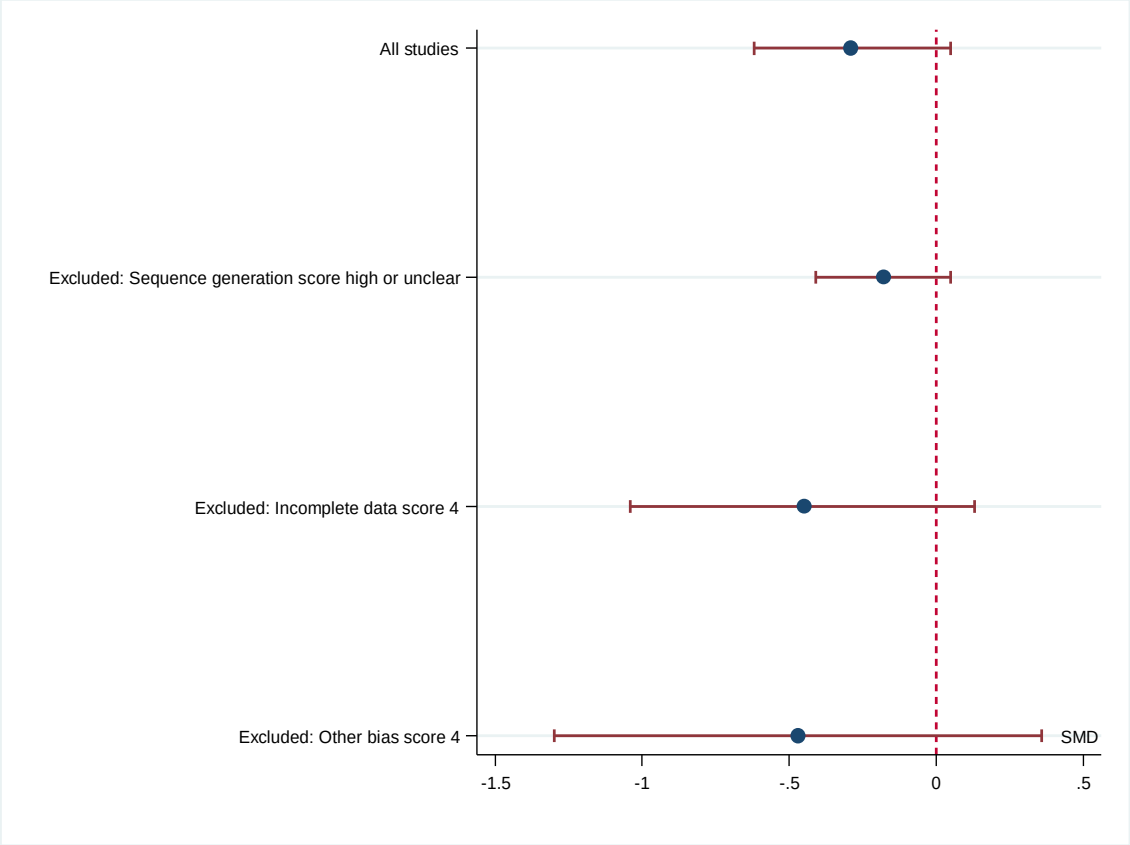
11.5 SENSITIVITY

11.5.1 Comparison: 12-step vs psychosocial intervention

11.5.1.1 Primary outcome: drug use

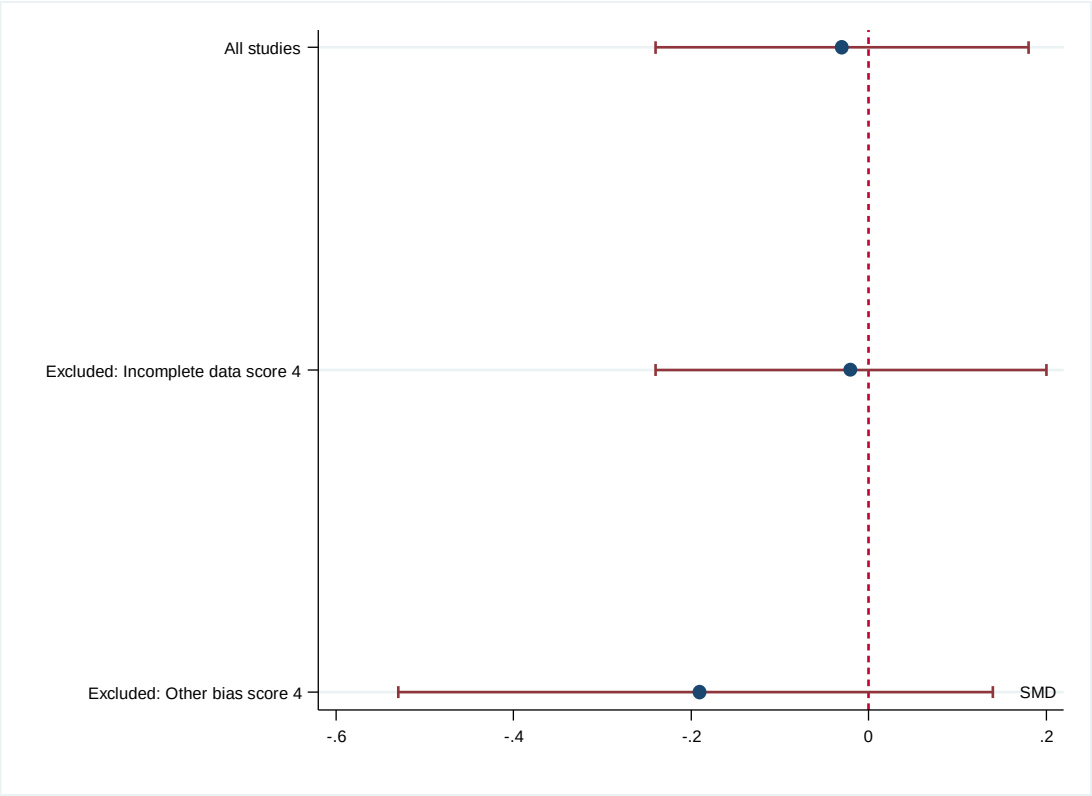
11.5.1.1.1 During treatment

Figure 3: Sensitivity Analysis. 12 step vs other psychosocial intervention, during treatment. Outcome: drug use. Sensitivity of effect size to excluding studies with high risk of bias on item: sequence generation (high or unclear), incomplete data (score of 4), and other bias (score of 4) respectively.



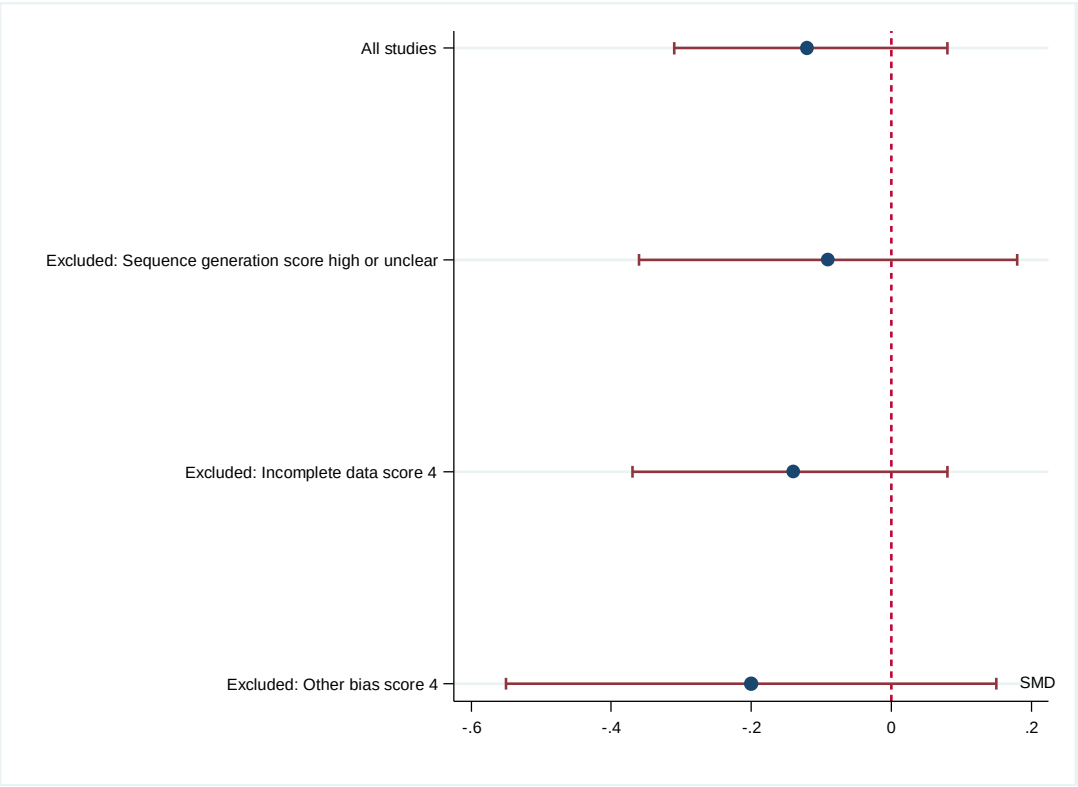
11.5.1.1.2 Post treatment

Figure 4: Sensitivity Analysis. 12 step vs other psychosocial intervention, post treatment. Outcome: drug use. Sensitivity of effect size to excluding studies with high risk of bias on item: incomplete data (score of 4), and other bias (score of 4) respectively.



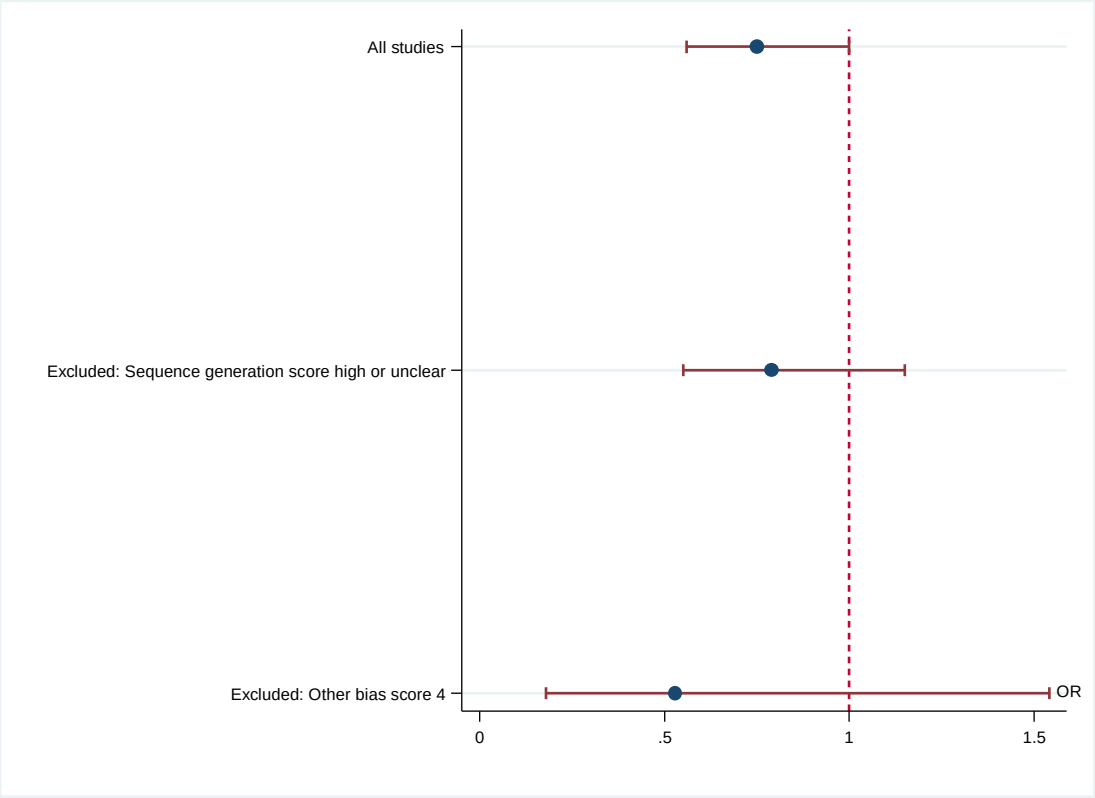
11.5.1.1.3 Follow-up

Figure 5: Sensitivity Analysis. 12 step vs other psychosocial intervention, post treatment. Outcome: drug use. Sensitivity of effect size to excluding studies with high risk of bias on item: incomplete data (score of 4), and other bias (score of 4) respectively.



11.5.1.2 Secondary outcome: retention

Figure 6: Sensitivity Analysis. 12 step vs other psychosocial intervention. Outcome: retention. Sensitivity of effect size to excluding studies with high risk of bias on item: sequence generation (high or unclear), and other bias (score of 4) respectively.



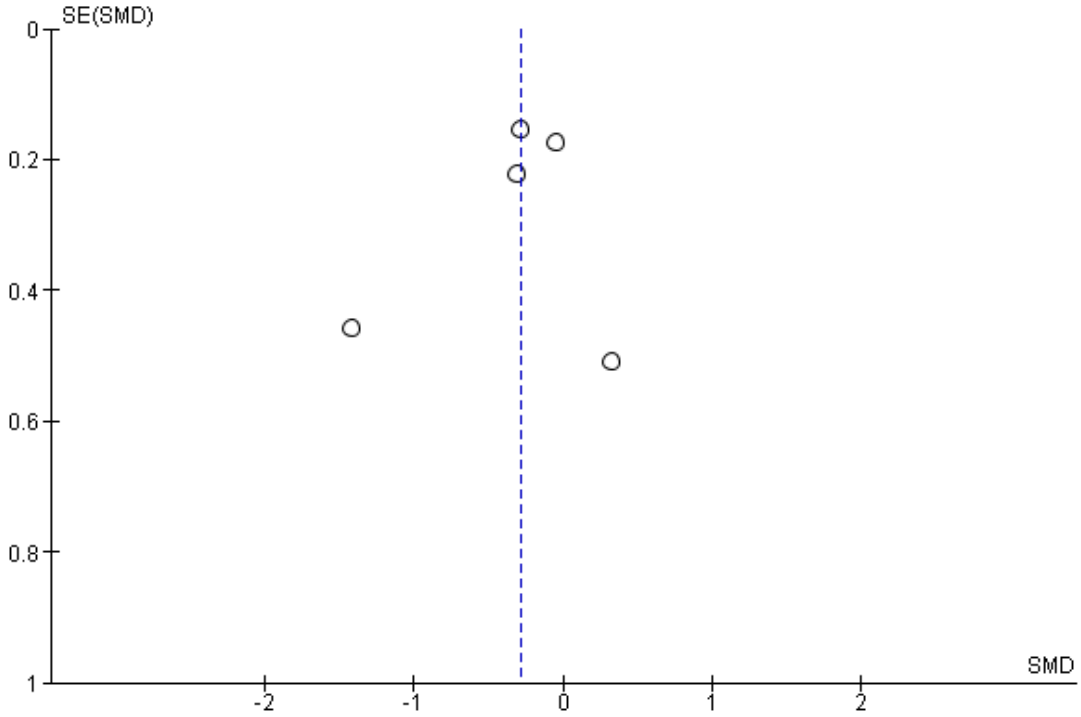
11.6 FUNNEL PLOTS

11.6.1 Comparison: 12-step vs psychosocial intervention

11.6.1.1 Primary outcome: drug use

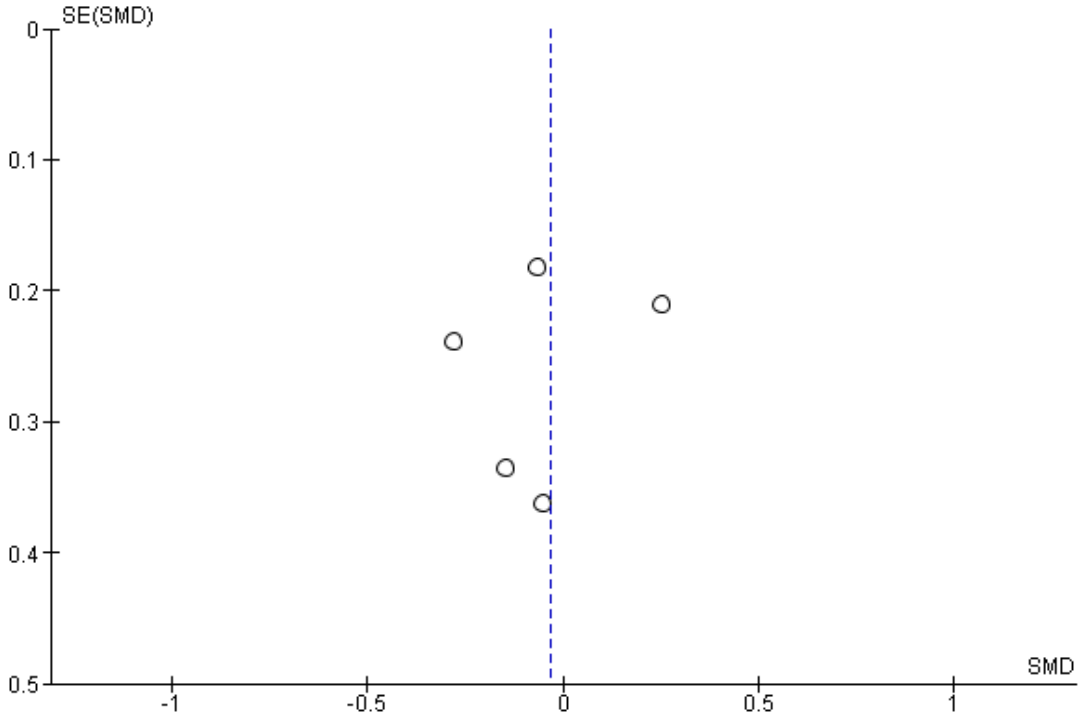
11.6.1.1.1 During treatment

Figure 7: Funnel plot for the comparison of 12-step vs other psychosocial interventions. Outcome: drug use. Time point: during treatment.



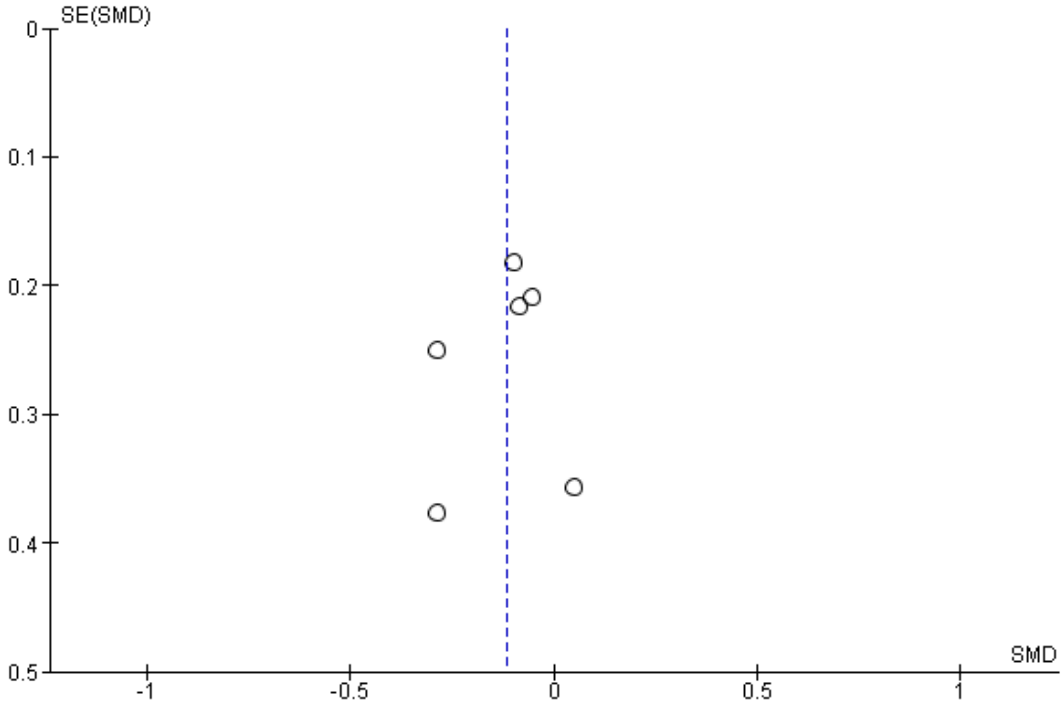
11.6.1.1.2 Post treatment

Figure 8: Funnel plot for the comparison of 12-step vs other psychosocial interventions. Outcome: drug use. Time point: post treatment.



11.6.1.1.3 Follow-up

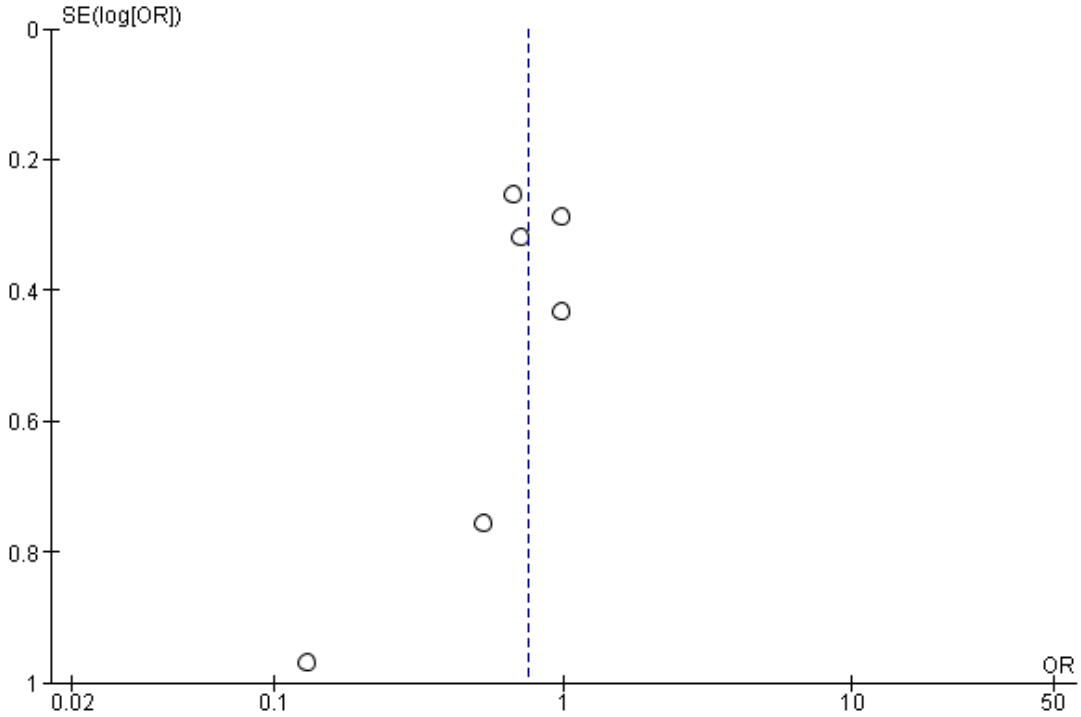
Figure 9: Funnel plot for the comparison of 12-step vs other psychosocial interventions. Outcome: drug use. Time point: follow-up.



11.6.1.2 Secondary outcome: retention

11.6.1.2.1 Post treatment

Figure 10: Funnel plot for the comparison of 12-step vs other psychosocial interventions. Outcome: retention. Time point: post treatment.



CampbellCollaboration

About this review

Illicit drug abuse is a globally recognised problem leading to high human, social and economic costs.

The 12-step program, modelled on the approach of Alcoholics Anonymous and adopted by Narcotics Anonymous and others, aims for complete abstinence. The 12-step approach is used both by self-help groups and for professional treatment called Twelve Step Facilitation (TSF). The broad applicability and low cost of the 12-step approach may appeal to policy makers.

This review examines the effectiveness of 12-step programs in reducing the use of illicit drugs. Secondary outcomes considered are on criminal behaviour, prostitution, psychiatric symptoms, social functioning, employment status, homelessness, and treatment retention.

The Campbell Collaboration
info@campbellcollaboration.org
Phone: (+47) 23 25 50 00

Mailing address:
P.O. Box 4404, Nydalen
N-0403 Oslo, Norway

Visiting address:
Pilestredet Park 7
(Entrance from Stensberggata)

Website:
www.campbellcollaboration.org